5-2020

# Low Latency Anomaly Detection with Imperfect Models

Samrat Nath
*University of Arkansas, Fayetteville*

Low Latency Anomaly Detection with Imperfect Models

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Engineering with a concentration in Electrical Engineering

by

Samrat Nath
Bangladesh University of Engineering and Technology
Bachelor of Science in Electrical and Electronics Engineering, 2014

May 2020
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

---

Jingxian Wu, Ph.D.
Dissertation Director

---

Roy A. McCann, Ph.D.
Committee Member

---

Haitao Liao, Ph.D.
Committee Member

---

Yue Zhao, Ph.D.
Committee Member

**Abstract**

The problem of anomaly detection deals with detecting abrupt changes/anomalies in the distribution of sequentially observed data in a stochastic system. This problem applies to many applications, such as signal processing, intrusion detection, quality control, medical diagnosis, etc. A low latency anomaly detection algorithm, which is based on the framework of quickest change detection (QCD), aims at minimizing the detection delay of anomalies in the sequentially observed data while ensuring satisfactory detection accuracy. Moreover, in many practical applications, complete knowledge of the post-change distribution model might not be available due to the unexpected nature of the change. Hence, the objective of this dissertation is to study low latency anomaly detection or QCD algorithms for systems with imperfect models such that any type of abnormality in the system can be detected as quickly as possible for reliable and secured system operations. This dissertation includes the theoretical foundations behind these low latency anomaly detection algorithms along with real-world applications. First, QCD algorithms are designed for detecting changes in systems with multiple post-change models under both Bayesian and non-Bayesian settings. Next, a QCD algorithm is studied for real-time detection of false data injection attacks in smart grids with dynamic models. Finally, a QCD algorithm for detecting wind turbine bearing faults is developed by analyzing the statistical behaviors of stator currents generated by the turbines. For all the proposed algorithms, analytical bounds of the system performance metrics are derived using asymptotic analysis and the simulation results show that the proposed algorithms outperform existing algorithms.

**Acknowledgments**

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Jingxian Wu for the continuous support of my Ph.D. study and research, for his patience, motivation, and immense knowledge. I thank him for teaching me every possible aspect of graduate-level research including how to approach positively towards difficult problems, how to present the research output to the research community effectively, and most importantly, how to do research. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Second, I am forever grateful to my family for their love, support, and belief in me that kept me motivated throughout the long journey of Ph.D. I cannot thank my parents enough for being so loving and caring, for selflessly dedicating their lives for me and my elder brother so that we can live better lives. I am also indebted to my elder brother, who has been my role model my whole life. Without his encouragement, I wouldn't have ever dreamed of pursuing a Ph.D degree.

Third, I take this opportunity to thank the support staff in the Department of Electrical Engineering, College of Engineering, and the office of Graduate School for assisting me with every query I ever had throughout my Ph.D. study. I would also like to thank my fellow labmates for the stimulating discussions and the fun we have had over the years, which made this journey much more enjoyable for me.

Last but not least, I am extremely grateful to my dissertation committee members, Dr. Roy A. McCann, Dr. Haitao Liao, and Dr. Yue Zhao. Their careful reviews and valuable comments have helped me significantly revise and improve the quality of the dissertation.

## Contents

## List of Figures

## List of Tables

## List of Papers

- **Chapter 2**, S. Nath and J. Wu, "Bayesian Quickest Change-point Detection with Multiple Candidates of Post-change Models", in *Proc. IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, U.S.A, 26-29 Nov 2018, pp. 51-55 (Published).

- **Chapter 3**, S. Nath and J. Wu, "Quickest Change Point Detection with Multiple Post-change Models" in *Journal of Sequential Analysis: Design Methods and Applications*, Jan. 2020 (Accepted).

- **Chapter 4**, S. Nath, I. Akingeneye, J. Wu, and Z. Han, "Quickest Detection of False Data Injection Attacks in Smart Grid with Dynamic Models", in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, Aug. 2019 (Accepted).

- **Chapter 5**, S. Nath, J. Wu, Y. Zhao, and W. Qiao, "Low Latency Bearing Fault Detection of Direct-drive Wind Turbines Using Stator Current", in *IEEE Access*, vol. 8, pp. 44163-44174, Mar. 2020 (Published).

# Chapter 1

# Introduction

## 1.1 Background and Motivation

Anomaly detection simply refers to the problem of detecting anomalies in sequentially observed data of a stochastic system. Anomaly detection is considered as an application of change detection problem, which is defined as the process of detecting the time instants at which the distribution of a stochastic process changes, i.e., an anomaly occurs in the system. The problem of anomaly/change detection arises in various branches of science and engineering. Change detection methods can be classified into two categories, offline and online change detection. In offline change detection, algorithms consider the entire data set at once, and estimate the change points retrospectively [1]. Usually, offline methods need to find the number of change points before estimating the location of each change point. On the contrary, online or real-time methods run simultaneously with the system process which is being monitored and use sequential analysis on currently observed samples to detect whether a change point has happened before the current time [2–7]. Online methods usually need to make tradeoff among various performance metrics such as average detection delay (ADD), probability of false alarm (PFA), and average run length (ARL) to false alarm, etc.

The problem of low latency anomaly detection can be studied under the quickest change detection (QCD) framework. QCD is an online change detection method, which aims to minimize the detection delay of a change point subject to constraints of an upper bound on PFA or a lower bound on ARL. The change point itself can be modeled as a random variable

with prior distributions. Based on the availability of information about the prior, change point detection methods can be classified into two categories, Bayesian and non-Bayesian (minimax) methods. If the prior probability of the change point is known, then Bayesian procedures, such as the well-known Shiryaev procedure [2], can be applied to minimize the ADD, under the constraints of an upper bound on PFA. It is shown in [3] that the Shiryaev procedure is asymptotically optimum when the PFA upper bound is small. When the prior probability of the change point is unknown, the non-Bayesian procedures, such as the cumulative sum (CUSUM) [4] method and the Shiryaev-Roberts (SR) procedure [5], perform detection under the minimax criterion, which aims at minimizing the expected delay with the worst-case change point distribution, under the constraint of a lower bound of ARL. The asymptotic optimality of the CUSUM and SR procedures are discussed in [6] and [7], respectively.

All the above-mentioned procedures require precise knowledge of the distribution functions of the stochastic process before and after the change point. In many applications, such as anomaly detection, it is relatively easier to obtain the prior-change distribution model from data collected through normal operation conditions. In contrast, it is usually difficult to obtain an accurate estimate of the post-change distribution model, especially for QCD where a decision needs to be made as soon as possible with a limited number of observations from the post-change distribution. In [8], a modified generalized likelihood ratio test (GLRT) is developed to take into consideration of some unknown parameters in the post-change distribution, and it is shown that the modified procedure can attain the same asymptotic lower bound of detection delay as the case of known post-change distribution. However, this algorithm is computationally very expensive, thus making it infeasible for

real-time implementation. In [9], a non-parametric quickest detection method that does not require prior knowledge of the distributions is proposed.

## 1.2 Applications

The topic of QCD has a wide range of applications in various fields, such as quality control, anomaly detection, intrusion detection [10], financial market analysis [11], medical diagnosis [12], climate change detection [13], speech and image analysis [14, 15] , etc.

Many practical problems related to power system applications can be modeled as the change detection problem, where real-time detection of any change in the system's operating conditions is very crucial. For example, in [16–18], the problem of detecting and/or identifying power system transmission line outages is studied by utilizing measurements from phasor measurement units (PMUs). Different types of wind-turbine bearing fault diagnosis methods have been developed in the literature using various techniques, such as wavelet, maximum likelihood, Markov process, artificial neural networks (ANN) and support vector machine (SVM), etc. [19–23]. False data injection attack (FDIA) in power systems is another example, where an adversary can launch cyber-attacks by compromising the measurements obtained by the supervisory control and data acquisition (SCADA) system or PMUs and thus cause significant damages to power grids. In [24–26], FDIA detection algorithms for power systems with dynamic models are studied.

Most of the above-mentioned works focus on the accuracy of change detection, and they seldom consider the detection delay. A small detection delay can result in timely remedial actions, thus prevent catastrophic results due to extended damages and reduce economic loss. Therefore, detection delay is an essential design parameter for QCD algorithm.

Moreover, some of the methods require extensive training before actual detection in order to estimate the unknown post-change parameters, and in practice, there might not be sufficient data for training purpose. For some applications, the post-change model might be from a finite set of possible models, that is, there are multiple hypotheses of the post-change models. For example, for the detection of wind turbine (WT) bearing fault, the fault could be caused by a finite number of defects, such as inner race fault, outer race fault, cage fault, and roller defect [27]. Also, in some cases, the post-change model might not have a finite support. For example, for the detection of false data injection attack in power system buses, the attack location and parameters are completely unknown [28]. Thus, there exists an impetus to develop efficient and robust online QCD algorithms to detect changes in systems with imperfect post-change models for ensuring reliable and secure operation.

### 1.3 Objectives

The goal of this dissertation is to develop new QCD algorithms that enhance the operation stability and security in various system applications where complete knowledge about the post-change regime is not available. This is achieved by detecting any kind of change in the system's normal operation mode as quickly as possible in real-time. These algorithms are developed in three folds.

First, QCD algorithms are designed in order to detect changes in systems with multiple post-change models under both Bayesian and non-Bayesian settings. The post-change model is considered to be from a finite set of possible models. In the Bayesian setting, the prior probability distribution of change point and the prior probabilities of post-change models are known, whereas in the non-Bayesian setting these probabilities are unknown. The algorithms

aim to minimize the ADD, subject to upper bounds on the PFA. The objectives are to derive theoretical bounds on ADD using asymptotic analysis and to compare the performance of proposed algorithms against other existing algorithms through simulation.

Second, a QCD algorithm is discussed with an objective of real-time detection of false data injection attacks in smart grids with dynamic models to improve the cyber-security. The algorithm aims to minimize the worst-case detection delays of cyber-attacks, subject to an upper bound of the false alarm rate. Unlike the previous two cases, here the post-change model i.e. post-attack model is not from a finite discrete set of models. Instead, the unknown post-change parameter has continuous support space. Furthermore, a dynamic state estimation algorithm needs to be developed in order to estimate and track the time-varying and non-stationary power grid states. The QCD algorithm is to be developed by analyzing the statistical properties of the results obtained from dynamic state estimations, such that the algorithm minimizes the detection delay while accurately distinguishing FDIA from sudden system changes.

Third, a QCD algorithm is developed with an objective of real-time detection of bearing faults of direct-drive wind turbines by analyzing the statistical behaviors of stator currents generated by the WT. In this case, the amplitude of stator current follows the Gamma distribution at a given frequency, and the presence of fault changes the parameters of the Gamma distribution. Since the signature of a fault can appear in one of the multiple possible frequencies, this problem also assumes post-change model to be from a finite set of possible models similarly as the first case. Based on the unique properties of WT bearing faults, the aim is to develop a new multi-candidate QCD algorithm that can combine the statistics of signals from multiple candidate frequencies.

These objectives exhibit scopes for improvement over existing algorithms in related works and provide a good foundation for future works.

## 1.4 Dissertation Outline

The outline of the rest of the dissertation is given as follows.

- *Chapter 2*: In this chapter, the topic of quickest change point detection for systems with multiple post-change models under Bayesian setting is studied. The post-change model is considered to be from a finite set of possible models. The objective is to minimize the ADD, subject to upper bounds on the PFA. With asymptotic analysis, the theoretical ADD of the proposed algorithm under Bayesian setting is expressed as closed-form functions of the PFA upper bound, prior probabilities of change points and models, and the Kullback-Leibler divergence among various models. The proposed algorithm is shown to be asymptotically optimum in terms of ADD. The algorithm is simulated in order to analyze their performances and compare them against other existing algorithms.

- *Chapter 3*: This chapter is the non-Bayesian counterpart of the topic studied in Chapter 2. The problem is formulated with the same objective of minimizing the ADD, subject to upper bounds on the PFA. Under non-Bayesian setting, the asymptotic theoretical ADD depends on the PFA upper bound, number of post-change models, mean of the prior distribution of change-point, and the Kullback-Leibler divergence among various models. The asymptotic analysis proves that the proposed algorithm is optimum in terms of ADD. Simulation results are illustrated to analyze and compare the performance of the proposed algorithm against other existing algorithms.

- *Chapter 4*: In this chapter, a quickest intrusion detection algorithm is studied in order

to detect FDIA in smart grids with time-varying dynamic models. The QCD algorithm aims at minimizing the worst-case detection delays of cyber-attacks, subject to an upper bound of the false alarm rate. Since power grid state transitions could be caused by either cyber-attacks or sudden change in loads or grid configurations, we propose to distinguish between FDIA and sudden system change by using a time-varying dynamic model, which can accurately capture the dynamic state transitions due to changes in system configurations. A dynamic state estimation algorithm is developed to estimate and track the time-varying and non-stationary power grid states. The quickest detection algorithm is developed by analyzing the statistical properties of dynamic state estimations, such that the algorithm minimizes the worst-case detection delay while accurately distinguishing FDIA from sudden system changes. A Markov-chain-based analytical model is used to identify the detector's parameter and quantify its performance. Simulation results are demonstrated in order to analyze the FDIA detection performance of the proposed algorithm.

- *Chapter 5*: This chapter builds on the theoretical background of Chapter 3 to study the quickest detection of bearing faults of direct-drive WTs, by analyzing the statistical behaviors of stator currents generated by the WT in real time. At a given frequency, the amplitude of stator current follows the Gamma distribution, and the presence of fault will affect the parameters of the Gamma distribution. Since the signature of a fault can appear in one of the multiple possible frequencies, the signals on each possible frequency candidate are monitored simultaneously. Based on the unique properties of WT bearing faults, a new multi-candidate quickest detection algorithm is developed that can combine the statistics of signals from multiple candidate frequencies. The new algorithm can perform online detection of various possible bearing faults without requiring a separate training phase. The theoretical

performance of the proposed algorithm is analytically identified in the form of upper bounds of the PFA and ADD. Numerical results are shown in order to analyze the fault detection performance of the proposed algorithm.

- *Chapter 6*: Concluding remarks are drawn in this chapter by summarizing the major contributions of this dissertation and future research directions.

## References


[1] B. Zhang, J. Geng, and L. Lai, "Multiple change-points estimation in linear regression models via sparse group lasso," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2209–2224, 2015.

[2] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 22–46, 1963.

[3] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory of Probability & Its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

[4] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.

[5] S. Roberts, "A comparison of some control chart procedures," *Technometrics*, vol. 8, no. 3, pp. 411–430, 1966.

[6] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.

[7] M. Pollak and A. G. Tartakovsky, "Optimality properties of the shiryaev-roberts procedure," *Statistica Sinica*, pp. 1729–1739, 2009.

[8] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, 1998.

[9] T. Banerjee, H. Firouzi, and A. O. Hero, "Non-parametric quickest change detection for large scale random matrices," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 146–150.

[10] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blažek, and H. Kim, "Detection of intrusions in information systems by sequential change-point methods," *Statistical methodology*, vol. 3, no. 3, pp. 252–293, 2006.

[11] A. N. Shiryaev, "Quickest detection problems in the technical analysis of the financial data," in *Mathematical Finance—Bachelier Congress 2000*. Springer, 2002, pp. 487–521.

[12] R. Malladi, G. P. Kalamangalam, and B. Aazhang, "Online bayesian change point detection algorithms for segmentation of epileptic activity," in *2013 Asilomar Conference on Signals, Systems and Computers*. IEEE, 2013, pp. 1833–1837.

[13] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, "A review and comparison of changepoint detection techniques for climate data," *Journal of applied meteorology and climatology*, vol. 46, no. 6, pp. 900–915, 2007.

[14] D. Rybach, C. Gollan, R. Schluter, and H. Ney, "Audio segmentation for speech recognition using segment features," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2009, pp. 4197–4200.

[15] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE transactions on image processing*, vol. 14, no. 3, pp. 294–307, 2005.

[16] Y. C. Chen, T. Banerjee, A. D. Domínguez-García, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, 2016.

[17] G. Rovatsos, X. Jiang, A. D. Domínguez-García, and V. V. Veeravalli, "Statistical power system line outage detection under transient dynamics," *IEEE Transactions on Signal Processing*, vol. 65, no. 11, pp. 2787–2797, 2017.

[18] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644–1652, 2008.

[19] X. Gong, W. Qiao, and W. Zhou, "Incipient bearing fault detection via wind generator stator current and wavelet filter," in *36th Annual Conference on IEEE Industrial Electronics Society (IECON)*. IEEE, 2010, pp. 2615–2620.

[20] I. S. Bozchalooi and M. Liang, "Parameter-free bearing fault detection based on maximum likelihood estimation and differentiation," *Measurement Science and Technology*, vol. 20, no. 6, p. 065102, 2009.

[21] H. Ocak and K. A. Loparo, "A new bearing fault detection and diagnosis scheme based on hidden markov modeling of vibration signals," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5. IEEE, 2001, pp. 3141–3144.

[22] Y. Wang and D. Infield, "Neural network modelling with autoregressive inputs for wind turbine condition monitoring," in *International Conference on Sustainable Power Generation and Supply (SUPERGEN)*. IET, 2012, pp. 1–6.

[23] M.-S. An, S.-J. Park, J.-S. Shin, H.-Y. Lim, and D.-S. Kang, "Implementation of automatic failure diagnosis for wind turbine monitoring system based on neural network," in *Multimedia and Ubiquitous Engineering*. Springer, 2013, pp. 1181–1188.

[24] Q. Yang, L. Chang, and W. Yu, "On false data injection attacks against kalman filtering in power system dynamic state estimation," *Security and Communication Networks*, vol. 9, no. 9, pp. 833–849, Jun. 2016.

[25] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting false data injection attacks in ac state estimation," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2476–2483, Sep. 2015.

[26] J. James, Y. Hou, and V. O. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, Jul. 2018.

[27] X. Gong and W. Qiao, "Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3419–3428, 2013.

[28] I. Akingeneye and J. Wu, "Low latency detection of sparse false data injections in smart grids," *IEEE Access*, vol. 6, pp. 58 564–58 573, Oct. 2018.

# Chapter 2

## Bayesian Quickest Change Point Detection with Multiple Post-change Models

Samrat Nath, Jingxian Wu

### 2.1 Abstract

We study the quickest change point detection for systems with multiple possible post-change models. A change point is the time instant at which the distribution of a random process changes. We consider the case that the post-change model is from a finite set of possible models. Under the Bayesian setting, the objective is to minimize the average detection delay (ADD), subject to upper bounds on the probability of false alarm (PFA). The proposed algorithm is a threshold-based sequential test. With asymptotic analysis, the analytical ADD of the proposed algorithm is expressed as closed-form functions of the PFA, prior probabilities of change points and models, and the Kullback-Leibler divergences among various models. The proposed algorithm is shown to be asymptotically optimum in terms of ADD when the PFA tends to zero.

### 2.2 Introduction

Change point detection is the process of detecting the time instants at which the distribution of a random process changes [1]. It has a wide spectrum of applications in various science and engineering fields, such as quality control, anomaly detection, and seismology, etc. [2]. In many applications, it is relatively easier to obtain the distribution model before the change point, which usually corresponds to normal system operations. The post-change model, on

the other hand, might not be readily available due to the unexpected nature of the change. This problem is exacerbated for quickest change detection (QCD), which aims at minimizing the detection delay with only a small amount of post-change data for training post-change models [3, 4]. For many applications, the post-change model might be from a finite set of possible models, that is, there are multiple hypotheses of the post-change models. For example, for the detection of wind turbine bearing fault, the fault could be caused by a finite number of defects, such as inner race fault, outer race fault, cage fault, and roller defect [5].

Quickest change point detection methods attempt to detect the change point in real time by sequentially updating a test statistics with recently observed data. They are usually designed based on the tradeoff among several metrics, such as average detection delay (ADD), probability of false alarm (PFA), false alarm rate (FAR), and average run length (ARL) to false alarms, etc. Existing sequential change point detection methods can be classified into two categories, Bayesian and non-Bayesian (Minimax) methods. If the prior probability of the change point is known, then Bayesian procedures, such as the well-known Shiryaev procedure [6], can be applied to minimize the ADD, under the constraint of an upper bound on the PFA. [7] showed that the Shiryaev procedure is asymptotically optimal when the PFA upper bound is small. When the prior probability of the change point is unknown, non-Bayesian procedures, such as the cumulative sum (CUSUM) [8] method and the Shiryaev-Roberts (SR) procedure [9], aim at minimizing the delay with the worst-case change point distribution, under the constraint of a lower bound of ARL. The asymptotic optimality of the CUSUM and SR procedures are discussed in [10, 11]. The problem of QCD is studied in different fields of science and technology. In [12], a Bayesian QCD is formulated over wireless fading channels with energy constraints as a partially observable Markov decision problem

(POMDP) and the optimal stopping rules are shown to have weak threshold structure. The QCD problem to detect a point of disruption in centralized multi-sensor network is studied in [13]. A QCD algorithm is proposed in order to detect false data injection attacks (FDIA) in smart grids with time-varying dynamic models in [14].

All the above procedures are developed for binary hypothesis testing, i.e. for systems with single pre-change and single post-change model, and they require precise knowledge about the distribution models before and after the change. There are limited works with unknown or uncertain post-change models. A Bayesian QCD algorithm for system with multiple candidates of post-change models is developed in [15]. The author in [16] proposed two methods for detecting a post-change distribution with an unknown parameter. In the first method, the detection is performed by using a mixture post-change distribution, which is obtained by averaging the set of possible post-change distributions with prior distributions of the unknown parameter. The second method is based on the generalized likelihood ratio test (GLRT), where the unknown parameter of the post-change model is estimated by maximizing the likelihood ratio. The GLRT-based method still requires a generous amount of post-change training data to tune the unknown parameter for the GLRT. A low-complexity adaptive-CUSUM method is presented in [14] for estimating unknown statistics of post-change distributions by using a normalized Rao test statistic [17].

The objective of this chapter is to design sequential quickest change detection algorithms for systems with multiple possible post-change models under Bayesian setting. The algorithms are developed to minimize the average detection delay (ADD), under the constraint on the upper bounds of the probability of false alarm (PFA). Under the Bayesian setting,

the algorithm is developed by analyzing the likelihood ratio of the change point, the computation of which relies on the prior probabilities of change point and prior probabilities of different post-change models. The performances of the proposed algorithms are analytically quantified in terms of exact or asymptotic bounds on PFA and ADD. It is shown that when the PFA is small, the proposed algorithms are asymptotically optimal in terms of ADD minimization under a certain PFA upper bound. Numerical results demonstrate that the proposed algorithms outperform existing algorithms in the literature.

The rest of this chapter is organized as follows. Section 2.3 presents the assumptions and problem formulation. The Bayesian detection algorithm, along with the corresponding theoretical analysis, are given in details in Section 2.4. Section 2.5 demonstrates the performance of the algorithms through numerical results, and Section 2.6 concludes the chapter.

## 2.3 Problem Formulation

Consider a sequentially observed random sequence, $X_n$, $n = 1, 2, \cdots$. Let $\mathcal{F}_n^X = \sigma(\mathbf{X}^{1:n})$ be the $\sigma$-algebra generated by $\mathbf{X}^{1:n}$. Assume there is an unknown change point $\theta$, such that the distributions of the random sequence are different before and after $\theta$. Denote the probability density function (PDF) of the random sequence before the change point as $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$ for $n < \theta$. The distribution after the change point could be one of a finite number of possible distribution models, denoted as $f_{i,n}(X_n|\mathbf{X}^{1:n-1})$, for $n \geq \theta$ and $i = 1, 2, \cdots, M$, with $M < \infty$. Denote the index of true post-change distribution as $\beta$, where $\beta \in \{1, \cdots, M\}$ is unknown.

In a Bayesian setting, the change point $\theta$ is random with prior probability mass function (PMF) $\mathbf{P}(\theta = k) = \pi_k$, for $k = 1, 2, \cdots$. The post-change model index is random with prior PMF $\mathbf{P}(\beta = i) = \omega_i$, for $i = 1, \cdots, M$.

Let $\mathbf{P}_{k,i}$ and $\mathbb{E}_{k,i}$ denote the probability measure and the corresponding expectation operator when the change occurs at $\theta = k < \infty$ and the post-change model index is $\beta = i$. Under $\mathbf{P}_{k,i}$, the conditional PDF of $X_n$ is $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$ for $n < k$, and it is $f_{i,n}(X_n|\mathbf{X}^{1:n-1})$ for $n \geq k$. For any $k < \infty$, we have $\mathbf{P}_k = \sum_{i=1}^{M} \omega_i \mathbf{P}_{k,i}$ and $\mathbb{E}_k = \sum_{i=1}^{M} \omega_i \mathbb{E}_{k,i}$. Denote $\mathbf{P}_\infty$ and $\mathbb{E}_\infty$ as the probability measure and expectation operator for the data sequence before the change point, that is, under $\mathbf{P}_\infty$, the conditional PDF of $X_n$ is $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$.

We need to design a test in order to detect the change point $\theta$ based on the sequentially observed data $X_n$. Denote $\hat{\theta}$ as the estimated value of $\theta$. A sequential test $\delta$ can be defined as a mapping from $\mathcal{F}_n^X$ to $\hat{\theta} \in \{1, \cdots, n\}$, such that $\delta(\mathcal{F}_n^X) = \hat{\theta}$. The test needs to be designed by optimizing with respect to two performance metrics, the PFA and ADD.

For a given test $\delta$, the PFA and ADD are defined, respectively, as

$$\text{PFA}(\delta) = \mathbf{P}(\hat{\theta} < \theta | \mathcal{F}_n^X) \tag{2.1}$$

$$\text{ADD}(\delta) = \mathbb{E}[\hat{\theta} - \theta | \hat{\theta} \geq \theta] \tag{2.2}$$

The objective is to minimize the ADD, subject to a constraint on the PFA. The problem can thus be formulated as

$$\begin{aligned} \text{(P1)} \qquad & \text{minimize} \qquad \text{ADD}(\delta) \\ & \text{subject to} \qquad \text{PFA}(\delta) < \alpha \end{aligned}$$

We propose the solution to this problem under Bayesian setting in the following section.

### 2.4 Quickest Change Detection Algorithm for Bayesian Setting

In this section, we develop the algorithm that can detect the change point with minimum delay under the Bayesian setting.

### 2.4.1 Detection Algorithm

At any moment $n$, the detector needs to make a decision between two hypotheses

$$\mathcal{H}_1 : \theta \leq n$$

$$\mathcal{H}_0 : \theta > n$$

Define the ratio of the posterior probabilities as

$$\Delta(n) = \frac{\mathbf{P}(\mathcal{H}_1|\mathcal{F}_n^X)}{\mathbf{P}(\mathcal{H}_0|\mathcal{F}_n^X)}. \tag{2.3}$$

Based on Bayes' rule, we have

$$\Delta(n) = \frac{\sum_{k=1}^{n} \pi_k \cdot d\mathbf{P}(\mathbf{x}^{1:n}|\theta = k)}{\Omega_n \cdot d\mathbf{P}(\mathbf{x}^{1:n}|\theta > n)} = \sum\nolimits_{i=1}^{M} \omega_i \sum\nolimits_{k=1}^{n} \frac{\pi_k}{\Omega_n} \prod\nolimits_{t=k}^{n} \frac{f_{i,t}(X_t|\mathbf{X}^{1:t-1})}{f_{0,t}(X_t|\mathbf{X}^{1:t-1})} \tag{2.4}$$

where, $\Omega_n = \mathbf{P}(\theta > n) = \sum_{k=n+1}^{\infty} \pi_k$.

Define

$$Z_i^{k:n} = \sum\nolimits_{t=k}^{n} \log \frac{f_{i,t}(X_t|\mathbf{X}^{1:t-1})}{f_{0,t}(X_t|\mathbf{X}^{1:t-1})} \tag{2.5}$$

and

$$\Delta_i(n) = \sum\nolimits_{k=1}^{n} \frac{\pi_k}{\Omega_n} \exp\left(Z_i^{k:n}\right). \tag{2.6}$$

Then $\Delta(n)$ defined in (2.3) can be written as

$$\Delta(n) = \sum\nolimits_{i=1}^{M} \omega_i \Delta_i(n) \tag{2.7}$$

With $\Delta(n)$ defined in (2.7), the proposed quickest change detection algorithm is a threshold-based sequential test given as follows.

*Definition 2.1*: (Bayesian Quickest Change Detection) For a given PFA upper bound $\alpha$, the change point is detected as

$$\delta_1 : \hat{\theta}_1 = \inf\left\{n \geq 1 : \Delta(n) \geq \frac{1-\alpha}{\alpha}\right\} \tag{2.8}$$

It should be noted that the proposed algorithm in (2.8) can be considered as an extension of the well-known Shiryaev procedure [6], which only considers the case of one known post-change model. We will show next that the above algorithm is asymptotically optimal with respect to (P1).

### 2.4.2 Probability of False Alarm

We first study the PFA of the detection procedure defined in Definition 2.1.

*Lemma 2.1*: For the quickest change detection algorithm in Definition 2.1, the probability of false alarm is upper bounded by $\alpha$.

*Proof:* Let $p(n) = \mathbf{P}(\mathcal{H}_1|\mathcal{F}_n^X) = \mathbf{P}(\theta \leq n|\mathcal{F}_n^X)$. From (2.3), we have $\Delta(n) = \dfrac{p_n}{1-p_n}$, or equivalently,

$$p(n) = \frac{\Delta(n)}{\Delta(n)+1} = 1 - \frac{1}{\Delta(n)+1} \tag{2.9}$$

It is apparent that $p_n$ is an increasing function in $\Delta(n)$. From (2.8), we have

$$\Delta(\hat{\theta}_1) \geq \frac{1-\alpha}{\alpha} \Longrightarrow p(\hat{\theta}_1) \geq 1-\alpha. \tag{2.10}$$

From (2.1) and the definition of $p(n)$, the PFA can be calculated as

$$\mathrm{PFA}(\delta_1) = \mathbf{P}(\hat{\theta}_1 < \theta|\mathcal{F}_n^X) = 1 - p(\hat{\theta}_1) \tag{2.11}$$

Combining (2.11) with (2.10) completes the proof. ■

### 2.4.3 Average Detection Delay

In the Bayesian setting, the ADD defined in (2.2) can be computed as follows

$$\mathrm{ADD}(\delta) = \frac{\mathbb{E}(\hat{\theta} - \theta)^+}{\mathbf{P}(\hat{\theta} \geq \theta)} = \frac{1}{\mathbf{P}(\hat{\theta} \geq \theta)} \sum\nolimits_{k=1}^{\infty} \pi_k \; \mathbf{P}_k(\hat{\theta} \geq k) \; \mathbb{E}_k(\hat{\theta} - k | \hat{\theta} \geq k) \tag{2.12}$$

where $x^+ = \max(0, x)$.

To facilitate the ADD analysis, it is assumed that $\frac{1}{n} Z_i^{k:k+n}$ almost surely converges in probability $\mathbf{P}_i$ to a positive finite number $D_i$ [7], that is,

$$\frac{1}{n} Z_i^{k:k+n-1} \xrightarrow[n\to\infty]{\mathbf{P}_i - a.s.} D_i \quad \forall \; k < \infty \tag{2.13}$$

In the case of identically and independently distributed (i.i.d.) data models, we have $f_{i,t}(X_t | \mathbf{X}^{1:t-1}) = f_i(X_t)$, and $D_i = \mathbb{E}\left[\log \frac{f_i(X)}{f_0(X)}\right]$ is the Kullback-Leibler (KL) divergence between $f_i(X)$ and $f_0(X)$. Furthermore, similar to many studies in the literature [1, 3, 4, 7], we assume that the change point follows geometric distribution with parameter $\rho \in (0, 1)$, i.e., the PMF of change point is $\pi_k = (1 - \rho)^{k-1} \rho$.

The following asymptotic notations are used in the analysis. Consider two continuous functions $f(x)$ and $g(x)$ where $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \infty$. We have the following notations.

$$f(x) \underset{x \to x_0}{\preceq} g(x) \iff \lim_{x \to x_0} \frac{f(x)}{g(x)} \leq 1 \tag{2.14}$$

If both $f(x) \underset{x \to x_0}{\preceq} g(x)$ and $g(x) \underset{x \to x_0}{\preceq} f(x)$, then the two functions are called asymptotically equivalent as $x \to x_0$, and it is denoted as

$$f(x) \underset{x \to x_0}{\asymp} g(x) \iff \lim_{x \to x_0} \frac{f(x)}{g(x)} = 1 \tag{2.15}$$

*Theorem 2.1*: Assume the condition (2.13) holds and $\pi_k = (1-\rho)^{k-1}\rho$. As the PFA upper bound $\alpha \to 0$, we have

$$\mathbb{E}_k[(\hat{\theta}_1 - k)^+] \underset{\alpha\to 0}{\preceq} \min_i \left[\frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i}{D_i + |\log(1-\rho)|}\right]$$

*Proof:* To facilitate analysis, we first introduce a new stopping time with respect to each individual post-change distribution model as follows

$$\hat{\theta}_{1,i} = \inf\left\{n \geq 1 : \omega_i \Delta_i(n) \geq \frac{1-\alpha}{\alpha}\right\} \tag{2.16}$$

From the definition of $\Delta(n)$ in (2.7), it is evident that $\Delta(n) \geq \omega_i\Delta_i(n)$ for all $i$. Consequently,

$$\hat{\theta}_1 \leq \min_{i=1,2,\ldots,M} \hat{\theta}_{1,i} \tag{2.17}$$

The stopping time in (2.16) can be alternatively represented by

$$\hat{\theta}_{1,i} = \inf\left\{n \geq 1 : \log\Delta_i(n) \geq \log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i\right\} \tag{2.18}$$

Based on the definition of $\Delta_i(n)$ in (2.6), it is easy to show that

$$\log\Delta_i(n) \geq Z_i^{k:n} + \log\left(\frac{\pi_k}{\Omega_n}\right) \triangleq V_i^{k:n} \tag{2.19}$$

Define a new stopping time

$$\zeta_{1,i} = \inf\left\{n \geq 1 : V_i^{k:n} \geq \log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i\right\} \tag{2.20}$$

From (2.18)-(2.20), it is apparent that $\hat{\theta}_{1,i} \leq \zeta_{1,i}$, thus

$$\hat{\theta}_1 \leq \min_{i=1,2,\ldots,M} \hat{\theta}_{1,i} \leq \min_{i=1,2,\ldots,M} \zeta_{1,i} \tag{2.21}$$

For geometric pirors, we have

$$\lim_{n\to\infty} \frac{1}{n}\log\left(\frac{\pi_k}{\Omega_{k+n-1}}\right) = |\log(1-\rho)|. \tag{2.22}$$

Combining (2.13) with (2.22) yields

$$\frac{1}{n}V_i^{k:k+n-1} \xrightarrow[n\to\infty]{\mathbf{P}_i-a.s.} D_i + |\log(1-\rho)| \triangleq q_i. \tag{2.23}$$

Define

$$T_k = \sup\left\{n \geq 1 : \left|\frac{1}{n}V_i^{k:k+n-1} - q_i\right| > \epsilon\right\}. \tag{2.24}$$

If $\zeta_{1,i} - k > T_k$, then from (2.24) we have

$$\left|\frac{1}{\zeta_{1,i}-k}V_i^{k:\zeta_{1,i}-1} - q_i\right| \leq \epsilon, \quad \text{if } \zeta_{1,i} - k > T_k \tag{2.25}$$

which implies

$$\zeta_{1,i} - k \leq \frac{V_i^{k:\zeta_{1,i}-1}}{q_i - \epsilon}, \quad \text{if } \zeta_{1,i} - k > T_k \tag{2.26}$$

Based on the definition of $\zeta_{1,i}$ in (2.20), we have

$$V_i^{k:\zeta_{1,i}-1} < \log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i \tag{2.27}$$

Combining (2.26) and (2.27) results in

$$\zeta_{1,i} - k \leq \frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i}{q_i - \epsilon}, \quad \text{if } \zeta_{1,i} - k > T_k \tag{2.28}$$

When $\alpha < 0.5$ and $\epsilon < q_i$, we always have $\frac{\log\left(\frac{1-\alpha}{\alpha}\right)-\log\omega_i}{q_i-\epsilon} > 0$. Therefore the following inequality is true for both $\zeta_{1,i} - k > T_k$ and $\zeta_{1,i} - k \leq T_k$

$$\zeta_{1,i} - k \leq \frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i}{q_i - \epsilon} + T_k, \quad \text{if } \alpha < 0.5 \text{ and } \epsilon < q_i \tag{2.29}$$

Given the convergence condition in (2.23), we have $\mathbb{E}(T_k) < \infty$. Since $\epsilon$ can be arbitrarily small, we can let $\epsilon \to 0$. Thus when $\alpha \to 0$,

$$\mathbb{E}[\zeta_{1,i} - k] \underset{\alpha\to 0}{\preceq} \frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i}{D_i + |\log(1-\rho)|}. \tag{2.30}$$

Since $\hat{\theta}_1$ is a lower bound of $\zeta_{1,i}$ as in (2.21), we have

$$\mathbb{E}[\hat{\theta}_1 - k] \underset{\alpha \to 0}{\preceq} \min_i \left[ \frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log \omega_i}{D_i + |\log(1-\rho)|} \right]. \tag{2.31}$$

When $\alpha \to 0$, the right hand side of (2.31) is always positive. Thus the inequality in (2.31) still holds if we replace $\hat{\theta}_1 - k$ by $(\hat{\theta}_1 - k)^+$. This completes the proof. ■

From the results in Theorem 2.1, it can be seen the prior probability $\omega_i$ and the constant $D_i$ plays an important role in determining the ADD upper bound. If $\omega_i$ is very small, that is, the $i$-th post-change model is very unlikely, the value of $-\log \omega_i$ will be very large, and it will not affect the delay upper bound because the minimum is performed over all $M$ post-change models. Similarly, if $D_i$ is very small, that is, the difference between the $i$-th post-change model and the pre-change model is small, then the minimum operator will exclude its impact on the delay upper bound. Consequently, the delay upper bound is dominated by the post-change models that have large $\omega_i$ and/or large $D_i$, that is, those models that are likely to appear and have a big difference with the pre-change model.

In addition to the asymptotic upper bound in Theorem 2.1, we also have the asymptotic lower bound for the detection delay.

*Theorem 2.2*: Assume the condition (2.13) holds and $\pi_k = (1-\rho)^{k-1}\rho$. As the PFA upper bound $\alpha \to 0$, we have

$$\mathbb{E}_k[(\hat{\theta}_1 - k)^+] \underset{\alpha \to 0}{\succeq} \min_i \left[ \frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log \omega_i}{D_i + |\log(1-\rho)|} \right]$$

*Proof:* To simplify notation, define

$$L_{i\alpha} = \frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log \omega_i}{D_i + |\log(1-\rho)|}, \tag{2.32}$$

$$L_\alpha = \min_i \; L_{i\alpha} \tag{2.33}$$

Also, define the following events and their probabilities,

$$\mathcal{C}_{i,k} : \{k \leq \hat{\theta}_1 \leq k + (1-\epsilon)L_{i\alpha}\} \text{ and } \gamma^{\epsilon}_{i,k}(\hat{\theta}_1) = \mathbf{P}_k\{\mathcal{C}_{i,k}\}$$

$$\mathcal{C}_{k} : \{k \leq \hat{\theta}_1 \leq k + (1-\epsilon)L_{\alpha}\} \text{ and } \gamma^{\epsilon}_{k}(\hat{\theta}_1) = \mathbf{P}_k\{\mathcal{C}_{k}\}$$

where $0 \leq \epsilon < 1$ is a constant.

Since $L_\alpha = \min_i \ L_{i\alpha}$, it can be easily shown that $\mathcal{C}_k = \cap_i \, \mathcal{C}_{i,k}$. Thus

$$\gamma^{\epsilon}_{k}(\hat{\theta}_1) = \mathbf{P}_k\{\mathcal{C}_{1,k} \cap \mathcal{C}_{2,k} \cap \ldots \cap \mathcal{C}_{M,k}\} \leq \min_i \gamma^{\epsilon}_{i,k}(\hat{\theta}_1) \tag{2.34}$$

where the last inequality follows from the fact that for any events $Y$ and $Z$, $\mathbf{P}(Y \cap Z) = \mathbf{P}(Y) \cdot \mathbf{P}(Z|Y) \leq \mathbf{P}(Y)$.

Given the convergence condition in (2.13), it can be proven that

$$\lim_{\alpha \to 0} \gamma^{\epsilon}_{i,k}(\hat{\theta}_1) = 0 \quad \forall \ i = 1, \ldots, M, \ 0 < \epsilon < 1 \text{ and } k \geq 1 \tag{2.35}$$

The above equation can be proved by following a similar procedure for the proof of equation (3.31) in [7, Lemma 2]. Combining (2.34) with (2.35) yields

$$\lim_{\alpha \to 0} \gamma^{\epsilon}_{k}(\hat{\theta}_1) = 0 \quad \forall \ i = 1, \ldots, M, \ 0 < \epsilon < 1 \text{ and } k \geq 1 \tag{2.36}$$

Based on the Chebyshev inequality, for any $0 \leq \epsilon < 1$, we have

$$\begin{aligned} \mathbb{E}_k[(\hat{\theta}_1 - k)^+] &\geq [(1-\epsilon)L_\alpha] \ \mathbf{P}_k\{(\hat{\theta}_1 - k)^+ \geq (1-\epsilon)L_\alpha\} \\ &= [(1-\epsilon)L_\alpha] \ \mathbf{P}_k\{(\hat{\theta}_1 - k) \geq (1-\epsilon)L_\alpha\} \end{aligned} \tag{2.37}$$

where the last equality is based on the fact that the event $\{X^+ \geq A\}$ is true if and only if $\{X \geq A\}$ is true for $A > 0$, and $(1-\epsilon)L_\alpha$ is positive when $0 \leq \epsilon < 1$.

It is also evident that,

$$\mathbf{P}_k\{\hat{\theta}_1 - k \geq (1-\epsilon)L_\alpha\} \geq \mathbf{P}_k\{\hat{\theta}_1 \geq k\} - \gamma_k^\epsilon(\hat{\theta}_1) \tag{2.38}$$

which is based on the fact that for two events $Y$ and $Z$, $\mathbf{P}(Y \cup Z) \leq \mathbf{P}(Y) + \mathbf{P}(Z)$.

Based on Lemma 2.1, the PFA is upper bounded by $\alpha$. Thus

$$\alpha \geq \text{PFA}(\hat{\theta}_1) = \sum\nolimits_{i=1}^{\infty} \pi_i \mathbf{P}_i(\hat{\theta}_1 < i) \geq \pi_k \mathbf{P}_k(\hat{\theta}_1 < k)$$

Thus $\mathbf{P}_k(\hat{\theta}_1 < k) \leq \pi_k^{-1}\alpha$, or equivalently,

$$\mathbf{P}_k\{\hat{\theta}_1 \geq k\} = 1 - \mathbf{P}_k\{\hat{\theta}_1 < k\} \geq 1 - \pi_k^{-1}\alpha \tag{2.39}$$

Combining (2.37), (2.38) and (2.39), we get,

$$\mathbb{E}_k[(\hat{\theta}_1 - k)^+] \geq [(1-\epsilon)L_\alpha]\ [1 - \pi_k^{-1}\alpha - \gamma_k^\epsilon(\hat{\theta}_1)].$$

Since $\epsilon$ can be arbitrarily small, we can let $\epsilon \to 0$, then

$$\lim_{\alpha \to 0} \frac{\mathbb{E}_k[(\hat{\theta}_1 - k)^+]}{L_\alpha} \geq \lim_{\alpha \to 0}[1 - \pi_k^{-1}\alpha - \gamma_k^\epsilon(\hat{\theta}_1)] = 1$$

where the last equality is based on the fact that $\gamma_k^\epsilon(\hat{\theta}_1) \to 0$ as $\alpha \to 0$ as given in (2.36). This completes the proof. ■

The asymptotic lower bound in Theorem 2.2 is the same as the asymptotic upper bound in Theorem 2.1. The asymptotic convergence between the lower bound and upper bound indicates that the detection method in Definition 2.1 is asymptotically optimal. That is, the algorithm can asymptotically achieve the minimum detection delay because the asymptotic lower bound is also the asymptotic upper bound.

*Theorem 2.3*: Assume the condition (2.13) holds and $\pi_k = (1-\rho)^{k-1}\rho$. As the PFA upper bound $\alpha \to 0$, the quickest change detection presented in Definition 2.1 is asymptotically optimal with respect to (P1). The asymptotic ADD is

$$\mathrm{ADD}(\delta_1) \underset{\alpha\to 0}{\asymp} \min_i \left[\frac{\log\left(\frac{1-\alpha}{\alpha}\right) - \log\omega_i}{D_i + |\log(1-\rho)|}\right].$$

*Proof:* The results can be directly obtained by combining Theorems 2.1 and 2.2. ■

## 2.5 Numerical Results

Numerical results are presented in this section to demonstrate the performance of the proposed change detection algorithm under Bayesian setting as described in Definition 2.1. The algorithm utilizes the prior probabilities of change points as well as the prior probabilities of post-change models. All simulation results are obtained by averaging over 10,000 Monte-Carlo trials. The change point follows a geometric distribution with $\rho = 0.1$ in all simulations.

In the first example, we consider $M = 2$ possible post-change models. The pre-change and post-change distributions are zero-mean Gaussian distributions with variance $\sigma_i^2$, that is, $f_i \sim \mathcal{N}(0, \sigma_i^2)$. We have $\sigma_0^2 = 1$ for the pre-change distribution, and we will consider different combinations of the post-change parameters $(\sigma_1^2, \sigma_2^2)$. Figure 2.1 shows the average detection delay, $\mathrm{ADD}(\delta_1)$, as a function of the PFA upper bound $\alpha$ under various combinations of $(\sigma_1^2, \sigma_2^2)$ and model prior probability $\omega_1$. When $(\sigma_1^2 = 0.5, \sigma_2^2 = 1.5)$, we have $D_1 = 0.0966$ and $D_2 = 0.0473$. When $(\sigma_1^2 = 0.8, \sigma_2^2 = 1.2)$, we have $D_1 = 0.0116$ and $D_2 = 0.0088$. Under all configurations, the asymptotic analytical ADDs have the same slopes as their simulation counterparts. Thus the asymptotic results provide very good predictions regarding the trend of the detection delay. The performance difference of the three cases becomes smaller as the

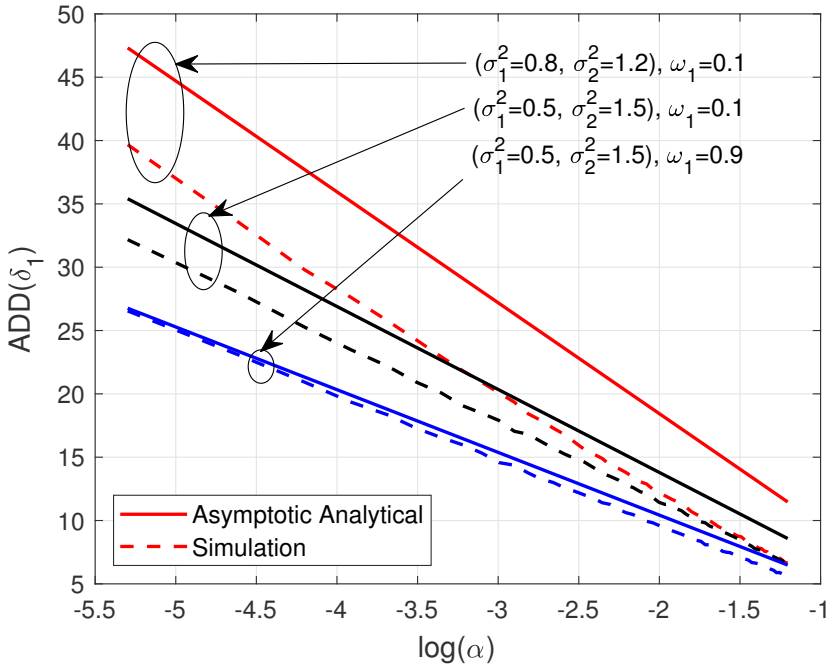


Figure 2.1: Average detection delay of the proposed Bayesian algorithm for various values of PFA upper bound $\alpha$.

PFA increases.

Figure 2.2 illustrates ADD($\delta_1$) as a function of the change point distribution parameter $\rho$ for the same model configurations used in Figure 2.1. The PFA upper bound is set as $\alpha = 10^{-3}$. The parameter $\rho$ is the inverse of the mean of the change point. A larger $\rho$ means a shorter average time period before the change point. The asymptotic analytical ADDs demonstrate the same trend as the simulation ADDs. As Theorem 2.3 suggests, the ADD decreases with the increase in $\rho$. That means if changes are more likely to occur at the beginning of a process, then the proposed Bayesian algorithm will show better performance in terms of ADD.

Figure 2.3 shows the PFA of the proposed Bayesian detection algorithm as a function of the PFA upper bound $\alpha$. There are $M = 2$ post-change models. The pre- and post-change

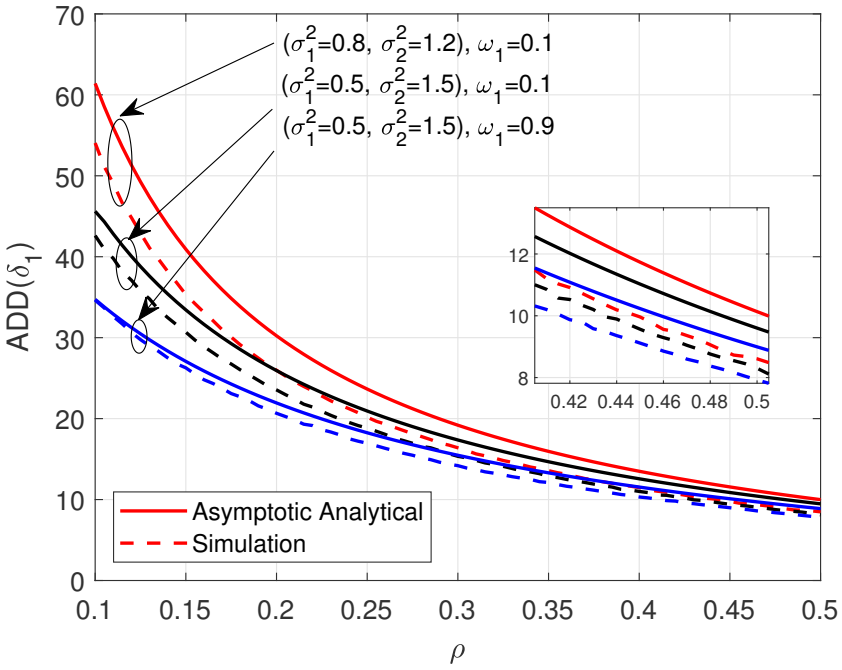


Figure 2.2: Average detection delay of the proposed Bayesian algorithm for various values of geometric distribution parameter $\rho$ with $\alpha = 10^{-3}$.

data follow exponential distributions with the parameters $\lambda_i$, for $i = 0, 1, 2$. We have $\lambda_0 = 1$ for the pre-change distribution, and $\lambda_1 = 0.5$ and $\lambda_2 = 1.5$ for the post-change distributions. It can be clearly observed that PFA obtained from numerical simulations is always below its upper bound as proved in Lemma 2.1. The analytical upperbound has the same trend as the simulated PFA under all system configurations.

Figure 2.4 compares the ADD of the proposed Bayesian algorithm with an adaptation of the well-known Shiryaev procedure. The adapted procedure exploits the mixture post-change distribution, which is obtained as $h(x) = \sum_i \omega_i f_i(x)$ [16]. The Shiryaev procedure is then employed by using $f_0$ and $h$ as the pre-change and post-change models, respectively. In this example, there are four post-change models, $f_i \sim \mathcal{N}(\mu_i, 1)$, with $\mu_1 = 0.6$, $\mu_2 = 0.8$, $\mu_3 = 1.2$, and $\mu_4 = 1.4$. The prior probabilities of these models are $\omega_1 = 0.1$, $\omega_2 = 0.2$, $\omega_3 = 0.3$, and

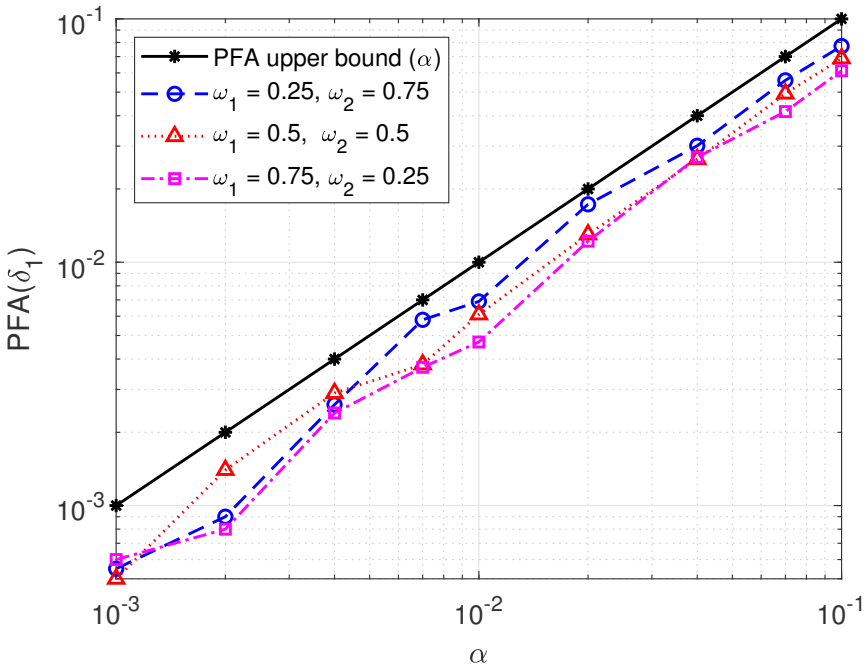


Figure 2.3: Probability of false alarm for the proposed Bayesian algorithm.

$\omega_4 = 0.4$, respectively. The pre-change model follows the distribution $f_0 \sim \mathcal{N}(1,1)$. The proposed algorithm outperforms the Shiryaev-Mixture algorithm under the entire range of PFA. At PFA $= 0.02$, the ADDs of the Shiryaev-Mixture and the proposed algorithms are 28 and 24, which corresponds to an improvement of 14% over the Shiryaev-Mixture algorithm.

## 2.6 Conclusion

We have proposed a threshold-based sequential test for Bayesian quickest change detection when there are multiple possible post-change models. The analytical ADD of the change point detection has been obtained through asymptotic analysis when the PFA is small. It has been shown that the proposed algorithm is asymptotically optimal in terms of average detection delay. Simulation results have shown that the asymptotic analytical results can predict the performance of the proposed algorithm, and the proposed algorithm outperforms

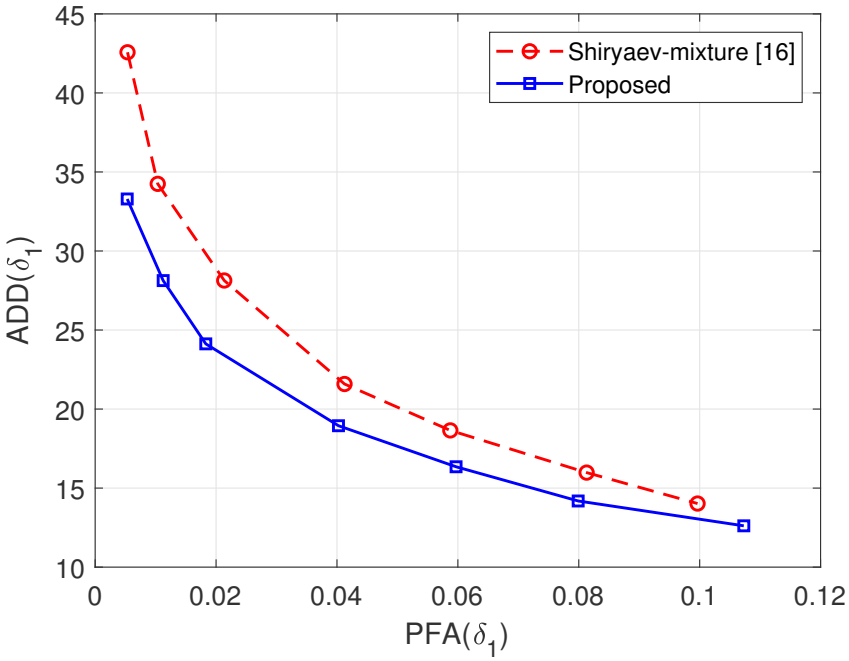


Figure 2.4: Comparison of the proposed Bayesian algorithm with adapted Shiryaev algorithms.

the Shiryaev procedure with a mixture post-change model.

## References


[1] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. Chapman and Hall/CRC, 2014.

[2] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowledge and information systems*, vol. 51, no. 2, pp. 339–367, 2017.

[3] H. V. Poor and O. Hadjiliadis, *Quickest detection*. Cambridge University Press Cambridge, 2009, vol. 40.

[4] V. V. Veeravalli and T. Banerjee, "Quickest change detection," in *Academic Press Library in Signal Processing*. Elsevier, 2014, vol. 3, pp. 209–255.

[5] X. Gong and W. Qiao, "Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3419–3428, 2013.

[6] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 22–46, 1963.

[7] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory of Probability & Its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

[8] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, Jun. 1954.

[9] S. Roberts, "A comparison of some control chart procedures," *Technometrics*, vol. 8, no. 3, pp. 411–430, 1966.

[10] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.

[11] M. Pollak and A. G. Tartakovsky, "Optimality properties of the shiryaev-roberts procedure," *Statistica Sinica*, pp. 1729–1739, 2009.

[12] X. Ren and L. Shi, "Bayesian quickest change detection under energy constraints over wireless sensor networks with correlated fading channels," in *11th IEEE International Conference on Control & Automation (ICCA)*, Taichung, Taiwan, Jun. 2014, pp. 1192–1197.

[13] V. Raghavan and V. V. Veeravalli, "Bayesian quickest change process detection," in *IEEE International Symposium on Information Theory (ISIT)*, Seoul, South Korea, Jun. 2009, pp. 644–648.

[14] S. Nath, I. Akingeneye, J. Wu, and Z. Han, "Quickest detection of false data injection attacks in smart grid with dynamic models," *IEEE Journal of Emerging and Selected Topics in Power Electronics (Accepted)*, Aug. 2019. [Online]. Available: https://doi.org/10.1109/JESTPE.2019.2936587

[15] S. Nath and J. Wu, "Bayesian quickest change point detection with multiple candidates of post-change models," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, USA, Nov. 2018.

[16] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[17] A. De Maio, "Rao test for adaptive detection in gaussian interference with unknown covariance matrix," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.

[18] A. G. Tartakovsky and V. V. Veeravalli, "Change-point detection in multichannel and distributed systems," *Applied Sequential Methodologies: Real-World Examples with Data Analysis*, vol. 173, pp. 339–370, 2004.

[19] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blažek, and H. Kim, "Detection of intrusions in information systems by sequential change-point methods," *Statistical Methodology*, vol. 3, no. 3, pp. 252–293, 2006.

[20] Y. Mei, "Efficient scalable schemes for monitoring a large number of data streams," *Biometrika*, vol. 97, no. 2, pp. 419–433, 2010.

[21] I. Akingeneye and J. Wu, "Low latency detection of sparse false data injections in smart grids," *IEEE Access*, vol. 6, pp. 58 564–58 573, Oct. 2018.

# Chapter 3

## Non-Bayesian Quickest Change Point Detection with Multiple Post-change Models

Samrat Nath, Jingxian Wu

### 3.1 Abstract

We study the sequential quickest change point detection for systems with multiple possible post-change models. A change point is the time instant at which the distribution of a random process changes. In many practical applications, the pre-change model can be easily obtained, yet the post-change distribution is unknown due to the unexpected nature of the change. In this chapter, we consider the case that the post-change model is from a finite set of possible models and consider the non-Bayesian setting, in which the prior probabilities of the change point and prior probabilities of possible post-change models are unknown. The objective is to minimize the average detection delay (ADD), subject to upper bounds on the probability of false alarm (PFA). Theoretical analysis is performed to quantify the analytical performance of the proposed algorithms in terms of exact or asymptotic bounds on PFA and ADD. It is shown through theoretical analysis that when PFA is small, both algorithms are asymptotically optimal in terms of ADD minimization for a given PFA upper bound. Numerical results demonstrate that the proposed algorithms outperform existing algorithms in the literature.

## 3.2 Introduction

The problem of change point detection deals with the process of detecting the time instants at which the distribution of a random process changes [1]. The applications of change point detection encompasses a wide spectrum in various science and engineering fields, such as quality control, anomaly detection, seismology, signal processing, biomedical engineering, etc. [2]. In case of many applications, it is relatively easier to obtain the model distribution before the change point, which usually corresponds to normal system operations. On the contrary, the post-change model might not be readily available due to the unknown nature of the change. This problem is exacerbated for quickest change detection (QCD), which aims at minimizing the detection delay with only a small amount of post-change data for training post-change models [3, 4]. For many applications, the post-change data might originate from a finite set of possible models, that is, there are multiple hypotheses of the post-change models. For example, for the detection of wind turbine bearing fault, the fault could be caused by a finite number of defects, such as inner race fault, outer race fault, cage fault, and roller defect [5].

QCD methods attempt to detect the change point in real time by sequentially updating a test statistics with recently observed data. They are usually designed based on the tradeoff among several metrics, such as average detection delay (ADD), probability of false alarm (PFA), false alarm rate (FAR), and average run length (ARL) to false alarms, etc. Based on the availability of the prior information regarding the change point, existing sequential change point detection methods can be classified into two categories, Bayesian and non-Bayesian (Minimax) methods. In the Bayesian methods, the prior probability of the change

point is known and utilized in designing the detector. Well-known Bayesian procedures, such as the Shiryaev procedure [6], can be applied to minimize the ADD, under the constraint of an upper bound on the PFA. The Shiryaev procedure is asymptotically optimal when the PFA upper bound is small [7]. On the other hand, for non-Bayesian procedures, the prior probability of the change point is unknown. Non-Bayesian procedures, such as the cumulative sum (CUSUM) [8] method and the Shiryaev-Roberts (SR) procedure [9], aim at minimizing the delay with the worst-case change point distribution, under the constraint of a lower bound of ARL. In [10,11], the asymptotic optimality of the CUSUM and SR procedures are discussed in details. The problem of QCD has been studied in different areas of science and technology. In [12], a Bayesian QCD is formulated over wireless fading channels with energy constraints as a partially observable Markov decision problem (POMDP) and the optimal stopping rules are shown to have weak threshold structure. The QCD problem to detect a point of disruption in centralized multi-sensor network is studied in [13].

All the above procedures are developed for binary hypothesis testing, i.e. for systems with single pre-change and single post-change model, and they require precise knowledge about the distribution models before and after the change. There have been some limited works related to non-Bayesian formulation of the QCD problem with unknown or uncertain or multiple post-change models. For example, [18] studied the non-Bayesian QCD problem in multi-channel and distributed network systems. A non-parametric sequential method is proposed using multichannel generalization of the CUSUM procedure for the detection of intrusions in information systems in [19]. None of the above works provide theoretical analysis to quantify the performance of the algorithms. [20] performed sequential change point detection based on the sum of local CUSUM statistics of each possible post-change

model. An orthogonal matching pursuit CUSUM (OMP-CUSUM) algorithm is proposed to detect false data attack in power grid systems with unknown post-attack parameters while minimizing the detection delay in [21].

The goal of this chapter is to develop sequential QCD algorithms for systems with multiple possible post-change models under non-Bayesian setting. The algorithms are designed to minimize the average detection delay (ADD), while keeping the probability of false alarm (PFA) under a low level. Under the non-Bayesian setting, the algorithm is developed by analyzing all the likelihood ratios of the change point corresponding to different post-change models. The resultant procedure happens to be the sum of local Shiryaev-Roberts (SR) statistics for each post-change model. The performances of the proposed algorithms are analytically quantified in terms of exact or asymptotic bounds on PFA and ADD. It is shown that when the PFA is small, the proposed algorithms are asymptotically optimal in terms of ADD minimization under a certain PFA upper bound. Numerical results illustrate that the proposed algorithms outperform other existing algorithms.

The rest of this chapter is organized as follows. Section 3.3 presents the assumptions and problem formulation. The non-Bayesian detection algorithm, along with the corresponding theoretical analysis, are given in details in Section 3.4. Section 3.5 demonstrates the performance of the algorithms through numerical results, and Section 3.6 concludes the chapter.

## 3.3 Problem Formulation

Consider a sequentially observed random sequence, $X_n$, $n = 1, 2, \cdots$. Let $\mathcal{F}_n^X = \sigma(\mathbf{X}^{1:n})$ be the $\sigma$-algebra generated by $\mathbf{X}^{1:n}$. Assume there is an unknown change point $\theta$, such that the distributions of the random sequence are different before and after $\theta$. Denote the probability

density function (PDF) of the random sequence before the change point as $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$ for $n < \theta$. The distribution after the change point could be one of a finite number of possible distribution models, denoted as $f_{i,n}(X_n|\mathbf{X}^{1:n-1})$, for $n \geq \theta$ and $i = 1, 2, \cdots, M$, with $M < \infty$. Denote the index of true post-change distribution as $\beta$, where $\beta \in \{1, \cdots, M\}$ is unknown.

The change point $\theta$ is random with prior probability mass function (PMF) $\mathbf{P}(\theta = k) = \pi_k$, for $k = 1, 2, \cdots$. The post-change model index is random with prior PMF $\mathbf{P}(\beta = i) = \omega_i$, for $i = 1, \cdots, M$. In the non-Bayesian setting, both these priors are unknown.

Let $\mathbf{P}_{k,i}$ and $\mathbb{E}_{k,i}$ denote the probability measure and the corresponding expectation operator when the change occurs at $\theta = k < \infty$ and the post-change model index is $\beta = i$. Under $\mathbf{P}_{k,i}$, the conditional PDF of $X_n$ is $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$ for $n < k$, and it is $f_{i,n}(X_n|\mathbf{X}^{1:n-1})$ for $n \geq k$. For any $k < \infty$, we have $\mathbf{P}_k = \sum_{i=1}^{M} \omega_i \mathbf{P}_{k,i}$ and $\mathbb{E}_k = \sum_{i=1}^{M} \omega_i \mathbb{E}_{k,i}$. Denote $\mathbf{P}_\infty$ and $\mathbb{E}_\infty$ as the probability measure and expectation operator for the data sequence before the change point, that is, under $\mathbf{P}_\infty$, the conditional PDF of $X_n$ is $f_{0,n}(X_n|\mathbf{X}^{1:n-1})$.

We need to design a test in order to detect the change point $\theta$ based on the sequentially observed data $X_n$. Denote $\hat{\theta}$ as the estimated value of $\theta$. A sequential test $\delta$ can be defined as a mapping from $\mathcal{F}_n^X$ to $\hat{\theta} \in \{1, \cdots, n\}$, such that $\delta(\mathcal{F}_n^X) = \hat{\theta}$. The test needs to be designed by optimizing with respect to two performance metrics, the PFA and ADD.

For a given test $\delta$, the PFA and ADD are defined, respectively, as

$$\text{PFA}(\delta) = \mathbf{P}(\hat{\theta} < \theta | \mathcal{F}_n^X) \tag{3.1}$$

$$\text{ADD}(\delta) = \mathbb{E}[\hat{\theta} - \theta | \hat{\theta} \geq \theta] \tag{3.2}$$

The objective is to minimize the ADD, subject to a constraint on the PFA. The problem

can thus be formulated as

$$\begin{aligned} \text{(P1)} \qquad & \text{minimize} \qquad \text{ADD}(\delta) \\ & \text{subject to} \qquad \text{PFA}(\delta) < \alpha \end{aligned}$$

We propose the solution to this problem under non-Bayesian setting in the following section.

## 3.4 Quickest Change Detection Algorithm for Non-Bayesian Setting

In this section, we develop the algorithm that can detect the change point with minimum delay under a non-Bayesian setting. Under the non-Bayesian setting, the prior probabilities of the change point, $\pi_k$, for $k = 1, 2, \cdots$ and the prior probabilities of the post-change model, $\omega_i$, for $i = 1, \cdots, M$, are all unknown.

### 3.4.1 Detection Algorithm

At any moment $n$, the detector needs to make a decision between two hypotheses

$$\mathcal{H}_1 : \theta \leq n$$

$$\mathcal{H}_0 : \theta > n$$

Define

$$Z_i^{k:n} = \sum\nolimits_{t=k}^{n} \log \frac{f_{i,t}(X_t|\mathbf{X}^{1:t-1})}{f_{0,t}(X_t|\mathbf{X}^{1:t-1})}, \tag{3.3}$$

and

$$\Lambda(n) = \sum\nolimits_{i=1}^{M} \Lambda_i(n) \tag{3.4}$$

where,

$$\Lambda_i(n) = \sum_{k=1}^{n} \exp\left(Z_i^{k:n}\right). \tag{3.5}$$

With $\Lambda(n)$ defined in (3.4), the proposed quickest change detection algorithm under non-Bayesian setting is a threshold-based sequential test given as follows.

*Definition 3.1*: (Non-Bayesian Quickest Change Detection) For a given PFA upper bound $\alpha$, the change point is detected as

$$\delta_2 : \hat{\theta}_2 = \inf\left\{n \geq 1 : \Lambda(n) \geq \frac{M\overline{\theta}}{\alpha}\right\} \tag{3.6}$$

where

$$\overline{\theta} = \sum_{k=1}^{\infty} k\pi_k \tag{3.7}$$

is the prior mean of the change point.

Note that the statistic $\Lambda_i(n)$ is the Shiryaev-Roberts (SR) statistic [9] for detecting a change corresponding to the $i$-th post-change model. The detection procedure $\delta_2$ is therefore an extension of the SR procedure adapted to detect changes in system with multiple post-change models [18]. The non-Bayesian detection algorithm in (3.6) has a similar flavor as the algorithm proposed in [20]. The test statistic in [20, Equation. (9)] is the sum of the CUSUM statistics for each post-change hypothesis. In (3.6), the test statistic is the sum of the SR statistics for each post-change hypothesis. Next, we will show that the proposed detection method in Definition 3.1 is asymptotically optimum with respect to (P1).

### 3.4.2 Probability of False Alarm

We first study the PFA of the detection procedure defined in Definition 3.1.

*Lemma 3.1*: For the quickest change detection algorithm in Definition 3.1, the probability of false alarm is upper bounded by $\alpha$.

*Proof:* The statistic $\Lambda_i(n)$ defined in (3.5) can be written in a recursive form as

$$\Lambda_i(n+1) = \lambda_i(n+1)\left[1 + \Lambda_i(n)\right] \tag{3.8}$$

where $\lambda_i(n)$ is the likelihood ratior (LR) of the $i$-th post-change model at time $n$

$$\lambda_i(n) = \frac{f_{i,n}(X_n|\mathbf{X}^{1:n-1})}{f_{0,n}(X_n|\mathbf{X}^{1:n-1})} \tag{3.9}$$

It is straightforward that $\mathbb{E}_\infty[\lambda_i(n)] = 1$, and from (3.8)

$$\mathbb{E}_\infty[\Lambda_i(n+1)|\mathcal{F}_n^X] = 1 + \Lambda_i(n). \tag{3.10}$$

Thus $\Lambda_i(n)$ is a submartingale with respect to the probability measure $\mathbf{P}_\infty$. In addition, since $\mathbb{E}[\Lambda_i(1)] = \mathbb{E}[\lambda_i(1)] = 1$, we have $\mathbb{E}_\infty[\Lambda_i(n)] = n$.

Combining the definition of $\Lambda(n)$ in (3.4) with (3.10), we get

$$\mathbb{E}_\infty\left[\Lambda(n+1)|\mathcal{F}_n^X\right] = M + \Lambda(n). \tag{3.11}$$

Thus $\Lambda(n)$ is also a submartingale with respect to $\mathbf{P}_\infty$ and $\mathbb{E}_\infty\left[\Lambda(n)\right] = Mn$. Using Doob's submartingale inequality, we get

$$\mathbf{P}_\infty\{\hat{\theta}_2 < n\} = \mathbf{P}_\infty\left\{\max_{1\le k\le n}\Lambda(k) \ge \frac{M\overline{\theta}}{\alpha}\right\} \le \frac{n\alpha}{\overline{\theta}}$$

Therefore,

$$\mathrm{PFA}(\delta_2) = \sum_{k=1}^{\infty}\pi_k\mathbf{P}_\infty\{\hat{\theta}_2 < k\} \le \sum_{k=1}^{\infty}\frac{\pi_k k\alpha}{\overline{\theta}} = \alpha.$$

■

### 3.4.3 False Alarm Rate

For a given test $\delta$, the FAR is defined as

$$\mathrm{FAR}(\delta) = \frac{1}{\mathbb{E}_\infty(\hat{\theta})} \tag{3.12}$$

where $\mathbb{E}_\infty(\hat{\theta})$ is known as the ARL to false alarm.

*Lemma 3.2*: For the quickest change detection algorithm in Definition 3.1, the false alarm rate is upper bounded by $\alpha/\overline{\theta}$.

*Proof:* From (3.11), it is obvious that $\Lambda(n) - Mn$ forms a Martingale with respect to $\mathbf{P}_\infty$.

If $\mathbb{E}_\infty[\hat{\theta}_2] = \infty$, then $\text{FAR}(\delta_2) = 0$ and it is bounded by $\alpha/\overline{\theta}$.

If $\mathbb{E}_\infty[\hat{\theta}_2] < \infty$, then based on the optional stopping theorem, we have

$$\mathbb{E}_\infty[\Lambda(\hat{\theta}_2) - M\hat{\theta}_2] = \mathbb{E}_\infty[\Lambda(1) - M] = 0. \tag{3.13}$$

Thus

$$\mathbb{E}_\infty[\hat{\theta}_2] = \frac{1}{M}\mathbb{E}_\infty[\Lambda(\hat{\theta}_2)] \tag{3.14}$$

Combining (3.14) with (3.6) yields

$$\mathbb{E}_\infty[\hat{\theta}_2] \geq \frac{\overline{\theta}}{\alpha} \tag{3.15}$$

which implies $\text{FAR}(\delta_2) \leq \alpha/\overline{\theta}$. ■

### 3.4.4 Average Detection Delay

The asymptotic upper and lower bounds for the ADD of the detection method in Definition 3.1 are derived respectively in a similar manner as in Subsection 2.4.3.

*Theorem 3.1*: Assuming the condition (2.13) holds, as the PFA upper bound $\alpha \to 0$, we have

$$\mathbb{E}_k[(\hat{\theta}_2 - k)^+] \underset{\alpha \to 0}{\preceq} \min_i \frac{\log\left(\frac{M\overline{\theta}}{\alpha}\right)}{D_i}$$

*Proof:* The proof follows a similar procedure as the proof of Theorem 2.1. First define a new stopping time with respect to each individual post-change distribution model as follows

$$\zeta_{2,i} = \inf\left\{n \geq 1 : Z_i^{k:n} \geq \log\left(\frac{M\bar{\theta}}{\alpha}\right)\right\} \tag{3.16}$$

From (3.4) and (3.5), it is straightforward that $\log \Lambda(n) \geq \log \Lambda_i(n) \geq Z_i^{k:n}$. Thus

$$\hat{\theta}_2 \leq \min_{i=1,2,\ldots,M} \zeta_{2,i}. \tag{3.17}$$

Given the fact that $\frac{1}{n} Z_i^{k:k+n-1}$ almost surely converges to $D_i$ in probability $\mathbf{P}_i$ as $n \to \infty$ as in (2.13), we can define

$$\Gamma_k = \sup\left\{n \geq 1 : \left|\frac{1}{n} Z_i^{k:k+n-1} - D_i\right| > \epsilon\right\}. \tag{3.18}$$

If $\zeta_{2,i} - k > \Gamma_k$, then from (3.18) we have

$$\left|\frac{1}{\zeta_{2,i} - k} Z_i^{k:\zeta_{2,i}-1} - D_i\right| \leq \epsilon, \quad \text{if } \zeta_{2,i} - k > \Gamma_k \tag{3.19}$$

which implies

$$\zeta_{2,i} - k \leq \frac{Z_i^{k:\zeta_{2,i}-1}}{D_i - \epsilon}, \quad \text{if } \zeta_{2,i} - k > \Gamma_k \tag{3.20}$$

From (3.16), we have

$$Z_i^{k:\zeta_{2,i}-1} < \log\left(\frac{M\bar{\theta}}{\alpha}\right) \tag{3.21}$$

Combining (3.20) and (3.21) results in

$$\zeta_{2,i} - k \leq \frac{\log\left(\frac{M\bar{\theta}}{\alpha}\right)}{D_i - \epsilon}, \quad \text{if } \zeta_{2,i} - k > T_k \tag{3.22}$$

When $\alpha < M\bar{\theta}$ and $\epsilon < D_i$, we always have $\frac{\log\left(\frac{M\bar{\theta}}{\alpha}\right)}{D_i - \epsilon} > 0$. Therefore the following inequality is true for both $\zeta_{2,i} - k > \Gamma_k$ and $\zeta_{2,i} - k \leq \Gamma_k$

$$\zeta_{2,i} - k \leq \frac{\log\left(\frac{M\bar{\theta}}{\alpha}\right)}{D_i - \epsilon} + \Gamma_k, \quad \text{if } \alpha < M\bar{\theta} \text{ and } \epsilon < D_i \tag{3.23}$$

Given the convergence condition in (2.13), we have $\mathbb{E}(\Gamma_k) < \infty$. Setting $\epsilon \to 0$ and $\alpha \to 0$, we have

$$\mathbb{E}[\zeta_{2,i} - k] \underset{\alpha \to 0}{\preceq} \frac{\log\left(\frac{M\overline{\theta}}{\alpha}\right)}{D_i}. \tag{3.24}$$

Since $\hat{\theta}_2$ is a lower bound of $\zeta_{2,i}$ as in (2.21), we have

$$\mathbb{E}[\hat{\theta}_2 - k] \underset{\alpha \to 0}{\preceq} \min_i \frac{\log\left(\frac{M\overline{\theta}}{\alpha}\right)}{D_i}. \tag{3.25}$$

When $\alpha \to 0$, the right hand side of (2.31) is always positive. Thus the inequality in (2.31) still holds if we replace $\hat{\theta}_2 - k$ by $(\hat{\theta}_2 - k)^+$. This completes the proof. ■

From the results in Theorem 3.1, it can be seen that the delay upper bound is dominated only by the post-change models that have large $D_i$, that is, those models that have a big difference with the pre-change model.

*Theorem 3.2*: Assuming the condition (2.13) holds, as the PFA upper bound $\alpha \to 0$, we have

$$\mathbb{E}_k[(\hat{\theta}_2 - k)^+] \underset{\alpha \to 0}{\succeq} \min_i \frac{\log\left(\frac{M\overline{\theta}}{\alpha}\right)}{D_i}$$

*Proof:* The proof follows a similar procedure as the proof of Theorem 2.2. First define

$$Y_{i\alpha} = \frac{\log\left(\frac{M\overline{\theta}}{\alpha}\right)}{D_i}, \quad \text{and } Y_\alpha = \min_i Y_{i\alpha} \tag{3.26}$$

Define the following events and their respective probabilities,

$$\mathcal{D}_{i,k} : \{k \le \hat{\theta}_2 \le k + (1-\epsilon)Y_{i\alpha}\} \text{ and } \phi^\epsilon_{i,k}(\hat{\theta}_2) = \mathbf{P}_k\{\mathcal{D}_{i,k}\}$$

$$\mathcal{D}_k : \{k \le \hat{\theta}_2 \le k + (1-\epsilon)Y_{\alpha}\} \text{ and } \phi^\epsilon_{k}(\hat{\theta}_2) = \mathbf{P}_k\{\mathcal{D}_{k}\}$$

where $0 \le \epsilon < 1$ is a constant.

Since $Y_\alpha = \min_i \ Y_{i\alpha}$, we have $\mathcal{D}_k = \cap_i \ \mathcal{D}_{i,k}$, which implies

$$\phi_k^\epsilon(\hat{\theta}_2) \leq \min_i \phi_{i,k}^\epsilon(\hat{\theta}_2). \tag{3.27}$$

Similar to (3.28), it can be shown that

$$\lim_{\alpha \to 0} \phi_{i,k}^\epsilon(\hat{\theta}_2) = 0, \ \forall i = 1, \ldots, M, \ 0 < \epsilon < 1 \text{ and } k \geq 1. \tag{3.28}$$

Combining (3.27) with (3.28) yields

$$\lim_{\alpha \to 0} \phi_k^\epsilon(\hat{\theta}_2) = 0, \ \forall i = 1, \ldots, M, \ 0 < \epsilon < 1 \text{ and } k \geq 1 \tag{3.29}$$

Based on the Chebyshev inequality, for any $0 \leq \epsilon < 1$, we have

$$\mathbb{E}_k[(\hat{\theta}_2 - k)^+] \geq [(1-\epsilon)Y_\alpha] \ \mathbf{P}_k\{(\hat{\theta}_2 - k) \geq (1-\epsilon)Y_\alpha\} \tag{3.30}$$

Since $\{\hat{\theta}_2 \geq k\} = \{\hat{\theta}_2 - k \geq (1-\epsilon)Y_\alpha\} \cup \mathcal{D}_k$, we have

$$\mathbf{P}_k\{\hat{\theta}_2 - k \geq (1-\epsilon)Y_\alpha\} \geq \mathbf{P}_k\{\hat{\theta}_2 \geq k\} - \phi_k^\epsilon(\hat{\theta}_2). \tag{3.31}$$

Based on Lemma 3.1, the PFA is upper bounded by $\alpha$. Thus

$$\alpha \geq \text{PFA}(\hat{\theta}_2) = \sum_{i=1}^{\infty} \pi_i \mathbf{P}_i(\hat{\theta}_2 < i) \geq \pi_k \mathbf{P}_k(\hat{\theta}_2 < k),$$

and consequently,

$$\mathbf{P}_k\{\hat{\theta}_2 \geq k\} = 1 - \mathbf{P}_k\{\hat{\theta}_2 < k\} \geq 1 - \pi_k^{-1}\alpha \tag{3.32}$$

Combining (3.30), (3.31), and (3.32), we get,

$$\mathbb{E}_k[(\hat{\theta}_2 - k)^+] \geq [(1-\epsilon)Y_\alpha] \ [1 - \pi_k^{-1}\alpha - \phi_k^\epsilon(\hat{\theta}_2)].$$

Since $\epsilon$ can be arbitrarily small, we can let $\epsilon \to 0$, then

$$\lim_{\alpha \to 0} \frac{\mathbb{E}_k[(\hat{\theta}_2 - k)^+]}{Y_\alpha} \geq 1$$

This completes the proof. ■

The asymptotic lower bound in Theorem 3.2 is the same as the asymptotic upper bound in Theorem 3.1. The asymptotic convergence between the lower bound and upper bound indicates that the detection method in Definition 3.1 is asymptotically optimal.

*Theorem 3.3*: Assuming the condition (2.13) holds, as the PFA upper bound $\alpha \to 0$, the quickest change detection presented in Definition 3.1 is asymptotically optimal with respect to (P1). The asymptotic ADD is

$$\mathrm{ADD}(\delta_2) \underset{\alpha\to 0}{\asymp} \min_i \frac{\log\left(\frac{M\overline{\theta}}{\alpha}\right)}{D_i}$$

*Proof:* The results can be directly obtained by combining Theorems 3.1 and 3.2. ■

### 3.5 Numerical Results

Numerical results are presented in this section to demonstrate the performance of the proposed change detection algorithm under non-Bayesian settings as described in Definition 3.1. All simulation results are obtained by averaging over 10,000 Monte-Carlo trials. The change point follows a geometric distribution with $\rho = 0.1$ in all simulations.

Next, we study the performance of the algorithm in Definition 3.1 under the non-Bayesian setting. In the first example, there are $M = 2$ equiprobable post-change models, i.e., $\omega_1 = \omega_2 = 0.5$. The pre-change and post-change distributions are zero-mean Gaussian distributions with variance $\sigma_i^2$. We have $\sigma_0^2 = 1$ for the pre-change distribution, and we will consider different combinations of the post-change parameters $(\sigma_1^2, \sigma_2^2)$. Figure 3.1 shows $\mathrm{ADD}(\delta_2)$ as a function of the PFA upper bound $\alpha$. When $(\sigma_1^2 = 0.6, \sigma_2^2 = 1.4)$, we have $D_1 = 0.0554$ and $D_2 = 0.0318$. When $(\sigma_1^2 = 0.55, \sigma_2^2 = 1.45)$, we have $D_1 = 0.0739$ and $D_2 = 0.0392$. When $(\sigma_1^2 = 0.5, \sigma_2^2 = 1.5)$, we have $D_1 = 0.0966$ and $D_2 = 0.0473$. Similar to the Bayesian case, the asymptotic analytical ADDs for non-Bayesian method have similar

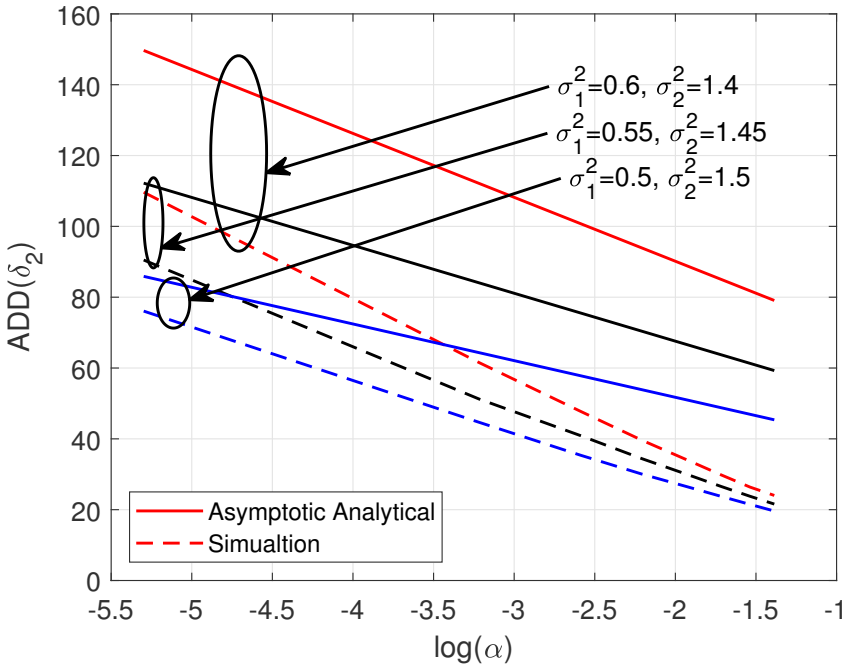


Figure 3.1: Average detection delay of the proposed non-Bayesian algorithm.

slopes as their simulation counterparts under different configurations. Thus the asymptotic results are very good predictors for the trend of the detection delay.

Figure 3.2 shows the PFA of the non-Bayesian algorithm as a function of the PFA upper bound $\alpha$ under different system configurations. There are $M = 2$ post-change models. The pre-change and post-change data follow Gaussian distributions with unit variance and mean $\mu_i$. We have $\mu_0 = 1$ for the pre-change distribution, and $\mu_1 = 0.5$ and $\mu_2 = 1.5$ for the post-change distributions. The simulated PFAs are always under the theoretical upper bounds as proved in Lemma 3.1. Under the non-Bayesian setting, the theoretical PFA upper bound is not as tight as its Bayesian counterpart. The PFA upper bound is about one order of magnitude higher than the results obtained from numerical simulations.

Figure 3.3 illustrates the performance of the non-Bayesian change detection algorithm described in Definition 3.1. In this example, there are $M = 3$ post-change models. The data

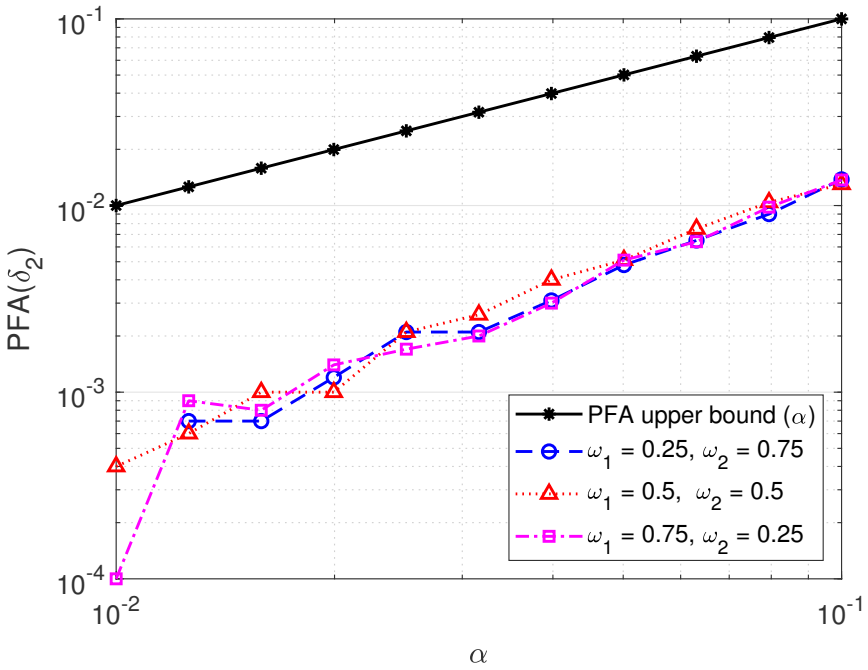


Figure 3.2: Probability of false alarm for the proposed non-Bayesian algorithm.

follow a two-dimensional multivariate Gaussian distribution with zero-mean and covariance matrix

$$R = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}.$$

The coefficient $r$ is set to 0 before the change point. After the change point, we set $r = 0.1, 0.5$, and 0.9 for the three post-change models, respectively. The prior probabilities of the post-change models are $\omega_1 = 0.1$, $\omega_2 = 0.3$, and $\omega_3 = 0.6$, respectively. Figure 3.3 compares the ADD of the proposed non-Bayesian algorithm with the ADD of three different non-Bayesian algorithms based on CUSUM procedure. The CUSUM-GLRT procedure [16] uses GLRT by estimating the unknown parameter, which corresponds to the post-change model in this example. The SUM-CUSUM procedure [20] exploits the sum of the local CUSUM statistics corresponding to the individual post-change models. The MAX-CUSUM

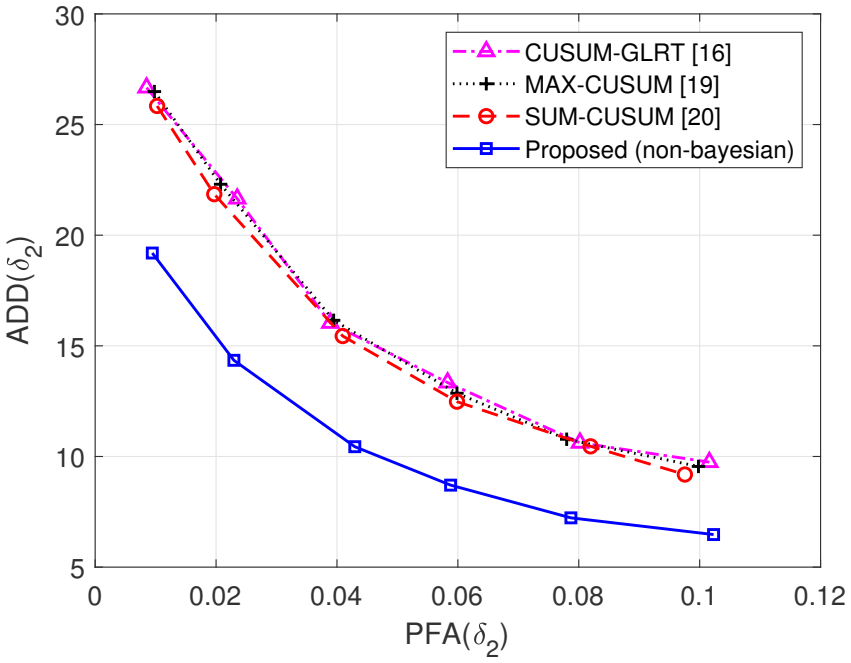


Figure 3.3: Comparison of the proposed non-Bayesian algorithm with adapted CUSUM algorithms.

procedure [19] uses the maximum of the local CUSUM statistics. The proposed non-Bayesian algorithm achieves significant performance gains over the existing CUSUM-based algorithms.

## 3.6 Conclusion

Quickest change point detection with multiple possible post-change models has been studied in this chapter. We have proposed two quickest change detection algorithms under the Bayesian and non-Bayesian settings, respectively. Theoretical analysis has been performed to obtain the PFA upper bounds and asymptotic bounds on ADD when the PFA is small. It has been shown that both algorithms are asymptotically optimal in terms of average detection delay. Numerical results have shown that the proposed algorithms outperform existing algorithms in terms of average detection delay under the same PFA constraints.

## References


[1] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. Chapman and Hall/CRC, 2014.

[2] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowledge and information systems*, vol. 51, no. 2, pp. 339–367, 2017.

[3] H. V. Poor and O. Hadjiliadis, *Quickest detection*. Cambridge University Press Cambridge, 2009, vol. 40.

[4] V. V. Veeravalli and T. Banerjee, "Quickest change detection," in *Academic Press Library in Signal Processing*. Elsevier, 2014, vol. 3, pp. 209–255.

[5] X. Gong and W. Qiao, "Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3419–3428, 2013.

[6] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 22–46, 1963.

[7] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory of Probability & Its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

[8] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, Jun. 1954.

[9] S. Roberts, "A comparison of some control chart procedures," *Technometrics*, vol. 8, no. 3, pp. 411–430, 1966.

[10] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.

[11] M. Pollak and A. G. Tartakovsky, "Optimality properties of the shiryaev-roberts procedure," *Statistica Sinica*, pp. 1729–1739, 2009.

[12] X. Ren and L. Shi, "Bayesian quickest change detection under energy constraints over wireless sensor networks with correlated fading channels," in *11th IEEE International Conference on Control & Automation (ICCA)*, Taichung, Taiwan, Jun. 2014, pp. 1192–1197.

[13] V. Raghavan and V. V. Veeravalli, "Bayesian quickest change process detection," in *IEEE International Symposium on Information Theory (ISIT)*, Seoul, South Korea, Jun. 2009, pp. 644–648.

[14] S. Nath, I. Akingeneye, J. Wu, and Z. Han, "Quickest detection of false data injection attacks in smart grid with dynamic models," *IEEE Journal of Emerging and Selected Topics in Power Electronics (Accepted)*, Aug. 2019. [Online]. Available: https://doi.org/10.1109/JESTPE.2019.2936587

[15] S. Nath and J. Wu, "Bayesian quickest change point detection with multiple candidates of post-change models," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, USA, Nov. 2018.

[16] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[17] A. De Maio, "Rao test for adaptive detection in gaussian interference with unknown covariance matrix," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.

[18] A. G. Tartakovsky and V. V. Veeravalli, "Change-point detection in multichannel and distributed systems," *Applied Sequential Methodologies: Real-World Examples with Data Analysis*, vol. 173, pp. 339–370, 2004.

[19] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blažek, and H. Kim, "Detection of intrusions in information systems by sequential change-point methods," *Statistical Methodology*, vol. 3, no. 3, pp. 252–293, 2006.

[20] Y. Mei, "Efficient scalable schemes for monitoring a large number of data streams," *Biometrika*, vol. 97, no. 2, pp. 419–433, 2010.

[21] I. Akingeneye and J. Wu, "Low latency detection of sparse false data injections in smart grids," *IEEE Access*, vol. 6, pp. 58 564–58 573, Oct. 2018.

# Chapter 4

## Quickest Detection of False Data Injection Attacks in Smart Grid with Dynamic Models

Samrat Nath, Israel Akingeneye, Jingxian Wu, Zhu Han

### 4.1 Abstract

A quickest intrusion detection algorithm is proposed to detect false data injection attacks (FDIA) in smart grids with time-varying dynamic models. The quickest detection algorithm aims at minimizing the worst-case detection delays of cyber-attacks, subject to an upper bound of the false alarm rate. Since power grid state transitions could be caused by either cyber-attacks or sudden change in loads or grid configurations, we propose to distinguish between FDIA and sudden system change by using a time-varying dynamic model, which can accurately capture the dynamic state transitions due to changes in system configurations. A dynamic state estimation algorithm is developed to estimate and track the time-varying and non-stationary power grid states. The quickest detection algorithm is developed by analyzing the statistical properties of dynamic state estimations, such that the algorithm minimizes the worst-case detection delay while accurately distinguishing FDIA from sudden system changes. A Markov-chain-based analytical model is used to identify the detector's parameter and quantify its performance. Simulation results demonstrate that the proposed algorithm can accurately detect and remove false data injections or system faults with minimum delays. The proposed algorithm can be implemented to harden intelligent electronic devices or supervisory control and data acquisition systems to improve their resilience to

cyber-attacks or system faults, thus improving the cyber-security of smart grids.

## 4.2 Introduction

A smart grid is a combination of electrical power infrastructure, smart meters, and a network of computers [1]. It uses information technologies to make intelligent decisions about the control and state of electrical power systems. Compared to conventional power grids, smart grid is more robust and efficient due to the advancement in system monitoring, energy management, and operation control. However, due to its dependence on cyber-infrastructure, a smart grid is prone to cyber-attacks [2]. Cyber-attacks can be performed by hacking into the communication network of smart grids, or by remotely accessing the remote terminal units (RTUs) installed at the substations [3]. For example, the supervisory control and data acquisition (SCADA) system of Iran's Natanz nuclear fuel-enrichment facility was attacked by a Stuxnet worm in July 2010 [4]. An adversary can launch cyber-attacks by compromising the measurement results obtained by the SCADA system or phasor measurement units (PMUs), such as the power injected into different buses or power flowing into the lines between the buses. False data injected in the measurement results will affect the real-time control of grid operations, thus cause significant damages to power grids. A comprehensive review of false data injection attack (FDIA) against modern power systems is given in [5]. To improve the cyber-security of smart grids, it is critical to ensure the integrity and confidentiality of the intelligent electronic devices (IEDs) in the network such as smart meters, RTUs, PMUs through hardware or software hardening [6, 7]. Tamper-proof hardware platforms can reduce avenues for FDIA.

A large number of algorithms have been developed to detect various forms of cyber-attacks in smart grids [8–13]. Most methods assume a static system model, where the system is in steady state and its measurements are quasi-static over time. However, in reality, the state of a power system varies with time due to the dynamic nature of system loads [14]. So, state estimation and FDIA detection algorithms require a dynamic model to track the time evolution of the system states, which can be utilized to detect and replace corrupted measurements in the system. A dynamic state estimator can capture the system transients due to sudden system changes in a faster and more accurate manner compared to its static counterpart. This is possible because of the dynamic state estimator's ability of using past state estimations to predict future state of the system one step ahead. A mismatch between newly collected measurements and their predicted values indicates that there have been sudden changes in the system such as loss of a large load, changes in network configurations, system faults, or malicious attacks that have modified some system measurements. It is vital to detect and identify these attacks as soon as possible in order to replace the corrupted measurements before they are processed by the state estimator.

Dynamic state estimation is important for the control and operations of a power grid [14–22]. Dynamic state estimation in many existing works is performed by using different versions of an extended Kalman filter (EKF) to filter predicted state variables [14–18]. In [21], FDIA is detected by tracking the dynamics of measurement variations in terms of the Kullback-Leibler divergence [23] between two probability distributions under normal and abnormal conditions. In [22], an online FDIA detection method is developed by analyzing temporally consecutive estimated system states using wavelet transform and deep neural network, which can effectively capture deviations in temporal data correlations of state vectors due to FDIA

scenarios. Most works utilize the estimation residual, which is the difference between the newly collected measurements and their corresponding predictions, to test the presence of FDIA. If the residual magnitude exceeds a certain threshold, a flag is raised indicating that either there is a sudden system change or FDIA. FDIA is distinguished from sudden system changes by analyzing correlated measurements in the location near the abnormality. In [14], if the measurements from neighboring buses fail the detection test simultaneously, a sudden change is declared. But, such a method might not be effective if false data are simultaneously injected into several neighboring buses with correlated measurements. This may lead to a mischaracterization of the attacks as sudden changes.

Most existing FDIA detection methods are developed to improve detection accuracy, with little or no attention given to detection delay. Detection delay is defined as the time difference between the launch and detection of a cyber-attack. Reducing detection delay is critical for improving cybersecurity [24]. A lower detection delay can shorten the response time so that remedial actions can be taken in a timely manner to significantly reduce the damages and economic losses caused by cyber-attacks. Detection delay can be reduced by employing algorithms from the quickest change detection (QCD) framework [25, 26], which aims at minimizing the average or worst-case detection delays while ensuring high detection accuracy. One of the most commonly used QCD procedure is the cumulative sum (CUSUM) procedure [25,27]. It has been shown in [28,29] that the CUSUM algorithm is asymptotically optimum, that is, it can asymptotically minimize the worst-case detection delay (WDD) when the false alarm rate goes to 0. However, implementation of CUSUM requires knowledge of the exact statistical distribution of the measurement under attack, which is usually unknown in practical applications [30]. An adaptive Rao-CUSUM test is proposed in [8] for false data

detection in smart grid, where the unknown distribution of data under attack is summarized by using the Rao test statistic [31]. In [24], an orthogonal matching pursuit CUSUM (OMP-CUSUM) algorithm is proposed to identify the buses under attack while minimizing the detection delay. Both [8] and [24] are developed under highly simplified linear static system models and they cannot capture the time-varying transient of power grids.

In this chapter, we develop a quickest intrusion detection algorithm for detecting FDIA in smart grids by using dynamic state estimations. This algorithm can be used to harden IEDs, PMUs, or SCADA system to improve their resilience to cyber-attacks or system faults. The detection method is designed to minimize the worst-case detection delay of FDIA subject to an upper bound of the false alarm rate, which is defined as the probability of falsely detecting an FDIA while the system is under normal operating conditions. One of the main challenges faced by FDIA detection is to distinguish power grid state changes caused by FDIA from those caused by a sudden system change, such as sudden load changes on certain buses. To address this challenge, we propose to use a locally linear but globally non-linear dynamic state model to represent the dynamic state transitions in power grids. The dynamic state evolution of the power grid is estimated and tracked by using an EKF-based dynamic state estimator, which estimates the current state by using both current measurements and predictions from past states. A sudden system change will affect the dynamic state transitions on all buses based on the physical model of the grid, and such state transitions can be accurately estimated by the dynamic state estimator based on SCADA or PMU measurements. On the contrary, FDIA or system faults might violate the dynamic state transitions determined by the model, and this may result in large residuals in estimation. Thus the employment of the dynamic state models can help distinguish FDIA from sudden

system change.

The quickest intrusion detection algorithm is developed by analyzing the statistical properties of the results obtained from dynamic state estimations. The problem is formulated as a hypothesis test performed on the residuals between the estimated and actual measurements. Since the false data attack vector is unknown at the detector, we propose a new normalized Rao-CUSUM test, which summarizes the unknown statistics of post-attack distributions by using a normalized Rao test statistic. Simulation results show that the normalization of Rao test statistic yields significantly lower FAR compared to un-normalized Rao test statistic under the same detection delay. The design parameter of the test is identified by using a Markov-chain based model of the test statistics through offline calculations. Once FDIA is detected, corrupted measurements are identified and replaced with their predicted values to ensure normal operations of the grid.

To summarize, this work has two main contributions. First, the detection algorithm aims at minimizing the worst-case detection delay of FDIA while ensuring high detection accuracy. The quickest detection algorithm is developed by using a new normalized Rao-CUSUM test that can accurately detect FDIA in a timely manner. Second, with a dynamic model and dynamic state estimations, the quickest detection algorithm can distinguish state transitions caused by FDIA from those caused by sudden system change, thus ensure the normal operations of the grid under both conditions.

The remainder of this chapter is organized as follows. Section 4.3 describes the system model and problem formulation. The dynamic model and dynamic state estimation are presented in Section 4.4. In Section 4.5, we develop the quickest detection algorithm by analyzing the statistical properties of the results from dynamic state estimations. In Section

Table 4.1: List of Notations in Chapter 4

| | |
|---|---|
| $m$ | total number of measurements |
| $\mathbf{z}_k$ | measurement vector at time instant $k$ |
| $\mathbf{x}_k$ | state vector |
| $\mathbf{R}_k$ | measurement error covariance matrix |
| $\mathbf{M}_k$ | prediction error covariance matrix |
| $\mathbf{h}(\mathbf{x}_k)$ | nonlinear function between $\mathbf{x}_k$ and $\mathbf{z}_k$ |
| $\mathbf{H}$ | Jacobian matrix |
| $\mathbf{a}$ | attack vector |
| $\mathbf{v}_k$ | residual vector |
| $\bar{\mathbf{v}}_k$ | whitened residual vector |
| $\mathbf{W}_k$ | whitening matrix |
| $L_k$ | log-likelihood ratio |
| $T_k$ | test statistic for false data detection |
| $A$ | threshold for detection rule |
| $\pi_j$ | steady-state probability of state $j$ |

4.6, a Markov-chain-based model is introduced to analytically evaluate the proposed false data detector. Simulation results are given in Section 4.7, and Section 4.8 concludes this chapter. A list of the notations used in this chapter is summarized in Table 4.1.

### 4.3 System Model

A power system with $N$ buses is considered. Without loss of generality, the first bus is assumed to be the reference. Define the set of buses connected to bus $i$ as $\mathcal{X}_i$ with cardinality $c_i = |\mathcal{X}_i|$. Denote the active and reactive power injections into bus $i$ as $P_i$ and $Q_i$, respectively. Similarly, the active and reactive power flows from bus $i$ to bus $j$ are denoted $P_{ij}$ and $Q_{ij}$, respectively, $\forall\ j \in \mathcal{X}_i$.

The power system collects measurements of both active and reactive power flows on different buses. The measurements are collected in such a way that the system becomes observable, i.e. all the state variables can be determined from the measurements. There

are many optimal approaches for sensor placement in order to make the system completely observable through collected measurements [32]. The power system provides a total of $m = m_1 + m_2 + 1$ measurements, where $m_1 = 2N$ is the number of active and reactive power injections, $m_2 = \sum_{i=1}^{N} |\mathcal{X}_i|$ is the number of active and reactive power flows. In addition to the power measurements, the measurement of the voltage magnitude at the reference bus is also available. Define the measurement vector as $\mathbf{z} = [z_1, z_2, \cdots, z_m]^T \in \mathcal{R}^{m\times 1}$, where $(\cdot)^T$ is the matrix transpose operator and $\mathcal{R}$ is the set of real numbers.

Define the state vector as $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathcal{R}^{n\times 1}$ for $n = 2N - 1$, where the first $N - 1$ elements of $\mathbf{x}$ are the voltage angles of $N - 1$ non-reference buses and the last $N$ elements are the voltage magnitudes of $N$ buses.

The relationship between the measurement vector $\mathbf{z}_k$ and the state vector $\mathbf{x}_k$, at an instant of time $k$ is expressed as

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{e}_k, \tag{4.1}$$

where $\mathbf{h}(\mathbf{x}_k) = [h_1(\mathbf{x}_k), \ldots, h_m(\mathbf{x}_k)]^T$ is a nonlinear function between the measurement vector $\mathbf{z}_k$ and the system state vector $\mathbf{x}_k$, and $\mathbf{e}_k \in \mathcal{R}^{m\times 1}$ is the measurement error vector at the sampling instant $k$. As shown in [14], we assume that the measurement noise $\mathbf{e}_k$ is zero-mean Gaussian distributed with covariance matrix $\mathbf{R}_k$.

Based on the observations in (4.1), the state estimator can obtain an estimate $\hat{\mathbf{x}}_k$ of the state variable $\mathbf{x}_k$. The state estimation results can be used to facilitate the detection of FDIA or system faults.

## 4.4 Dynamic State Estimation

In this section, we present a dynamic state estimation algorithm, which relies on previous estimates to predict future states of the system. The predicted states can, in turn, be used by the system operator for timely anomaly detection and other control decisions such as economic dispatch.

Consider the following state transition model, which describes the time behavior of the state vector, as

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k + \mathbf{w}_k, \tag{4.2}$$

where $\mathbf{F}_k \in \mathcal{R}^{n \times n}$ is a non-zero diagonal matrix, $\mathbf{G}_k \in \mathcal{R}^{n \times 1}$ is a non-zero column vector, and $\mathbf{w}_k \in \mathcal{R}^{n \times 1}$ is a white Gaussian noise vector with 0 mean and covariance matrix $\mathbf{Q}_k$.

The parameters $\mathbf{F}_k$ and $\mathbf{G}_k$ can be identified according to the Holt's exponential smoothing method [14]. The Holt's method performs smoothing over an original time series with two smoothing parameters, $\alpha$ and $\beta$, with values between 0 and 1. Denote the predicted state vector at time $k$ as $\tilde{\mathbf{x}}_k$. The Holt's method is expressed as

$$\tilde{\mathbf{x}}_{k+1} = \mathbf{a}'_k + \mathbf{b}'_k, \tag{4.3}$$

$$\mathbf{a}'_k = \alpha \mathbf{x}_k + (1 - \alpha)\tilde{\mathbf{x}}_k, \tag{4.4}$$

$$\mathbf{b}'_k = \beta \left[\mathbf{a}'_k - \mathbf{a}'_{k-1}\right] + (1 - \beta)\mathbf{b}'_{k-1}. \tag{4.5}$$

Combining (4.3)-(4.5) yields

$$\tilde{\mathbf{x}}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k, \tag{4.6}$$

where

$$\mathbf{F}_k = \alpha(1+\beta)\mathbf{I}_n,$$

$$\mathbf{G}_k = (1+\beta)(1-\alpha)\tilde{\mathbf{x}}_k - \beta\mathbf{a}'_{k-1} + (1-\beta)\mathbf{b}'_{k-1}.$$

The time-varying linear dynamic model in (4.2) can then be obtained by adding a zero mean Gaussian noise $\mathbf{w}_k$ to (4.6) to account for model uncertainties.

The proposed dynamic state estimator contains two steps: state forecasting and state estimation. Details are given as follows.

### 4.4.1 State Forecasting

One main advantage of the dynamic state estimator is its ability to use past state estimates to predict future system states. Let $\hat{\mathbf{x}}_k$ be the estimated state vector at time $k$ and $\mathbf{\Sigma}_k$ its error covariance matrix. The predicted state vector $\tilde{\mathbf{x}}_{k+1}$ and its error covariance matrix $\mathbf{M}_{k+1}$ at time $k$ can be obtained by performing the conditional expectation on (4.2) as follows

$$\tilde{\mathbf{x}}_{k+1} = \mathbb{E}\left[\mathbf{x}_{k+1} \mid \mathbf{x}_k = \hat{\mathbf{x}}_k\right] = \mathbf{F}_k\hat{\mathbf{x}}_k + \mathbf{G}_k, \tag{4.7}$$

$$\begin{aligned}\mathbf{M}_{k+1} &= \mathbb{E}\left[\left(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}\right)\left(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}\right)^T \mid \mathbf{x}_k = \hat{\mathbf{x}}_k\right] \\ &= \mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k + \mathbf{Q}_k,\end{aligned} \tag{4.8}$$

where $\mathbb{E}[.]$ is the expectation operator.

### 4.4.2 State Estimation

The state estimation, also known as state filtering, seeks to estimate the state at time $k+1$ by using both the predicated state vector, $\tilde{\mathbf{x}}_{k+1}$, obtained at the preceding step $k$, and the newly received measurement vector $\mathbf{z}_{k+1}$ at time $k+1$. During this stage, a new estimate

$\hat{\mathbf{x}}_{k+1}$ along with its error covariance matrix $\mathbf{\Sigma}_{k+1}$ are obtained at time $k+1$ by minimizing the objective function

$$J\left(\mathbf{x}_{k+1}\right) = \frac{1}{2}\left[\mathbf{z}_{k+1} - \mathbf{h}(\mathbf{x}_{k+1})\right]^T \mathbf{R}_{k+1}^{-1}\left[\mathbf{z}_{k+1} - \mathbf{h}(\mathbf{x}_{k+1})\right] + \frac{1}{2}\left[\left(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}\right)^T \mathbf{M}_{k+1}^{-1}\left(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}\right)\right]. \tag{4.9}$$

The estimate $\hat{\mathbf{x}}_{k+1}$ that minimizes the objective function in (4.9) can be obtained through an iterative extended Kalman filter (EKF) [14] as

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{\Sigma}^{(i)}\{\mathbf{H}^T(\mathbf{x}^{(i)})\mathbf{R}^{-1}[\mathbf{z} - \mathbf{h}(\mathbf{x}^{(i)})] - \mathbf{M}^{-1}[\mathbf{x}^{(i)} - \tilde{\mathbf{x}}]\}, \tag{4.10}$$

where $i$ denotes the iteration counter, $\mathbf{H}(\mathbf{x}) = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$ is the Jacobian matrix, and

$$\mathbf{\Sigma}^{(i)} = \left[\mathbf{H}^T(\mathbf{x}^{(i)})\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}^{(i)}) + \mathbf{M}^{-1}\right]^{-1}. \tag{4.11}$$

It should be noted that the subscript $k+1$ was omitted in (4.10) and (4.11) for simplicity. The proof for (4.10) is given below.

*Proof:*

The point $\mathbf{x}$, which minimizes (4.9) can be obtained by calculating the first derivative of $J(\mathbf{x})$ and setting it to zero. Define the first derivative of $J(\mathbf{x})$ as

$$\mathbf{g}(\mathbf{x}) = \frac{\partial \mathbf{J}\left(\mathbf{x}\right)}{\partial \mathbf{x}} = -\frac{\partial \mathbf{h^T}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{R^{-1}}\left[\mathbf{z} - \mathbf{h}(\mathbf{x})\right] + \mathbf{M^{-1}}\left(\mathbf{x} - \tilde{\mathbf{x}}\right). \tag{4.12}$$

The minimum point $\hat{\mathbf{x}}$ of $J(\mathbf{x})$ is calculated by solving

$$\mathbf{g}(\hat{\mathbf{x}}) = \mathbf{0}. \tag{4.13}$$

Given the non-linearity of (4.12), (4.13) is solved by iterative methods such as the Newton-Raphson method [33].

The Taylor series expansion of $\mathbf{g}(\mathbf{x})$ for $\mathbf{x} = \mathbf{x}^{(0)} + \Delta_{\mathbf{x}}$ is

$$\mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}^{(\mathbf{0})}) + \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} \mid_{\mathbf{x}=\mathbf{x}^{(0)}} \Delta_{\mathbf{x}}, \tag{4.14}$$

where $\mathbf{x}^{(0)}$ is the initial point and

$$\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{g}^{'}(\mathbf{x}) = \mathbf{H}^{T}(\mathbf{x})\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}) + \mathbf{M}^{-1}, \tag{4.15}$$

where $\mathbf{H}(\mathbf{x}) = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$.

According to the Newton-Raphson method [33], by setting (4.14) to zero, the increment $\Delta_{\mathbf{x}}$ is obtained as

$$\Delta_{\mathbf{x}} = -\left[\mathbf{g}^{'}(\mathbf{x}^{(0)})\right]^{-1} \mathbf{g}(\mathbf{x}^{(\mathbf{0})}). \tag{4.16}$$

Thus,

$$\mathbf{x} = \mathbf{x}^{(0)} - \mathbf{\Sigma}^{(0)}\mathbf{g}(\mathbf{x}^{(\mathbf{0})}), \tag{4.17}$$

where

$$\mathbf{\Sigma}^{(0)} = \left[\mathbf{g}^{'}(\mathbf{x}^{(0)})\right]^{-1} = \left[\mathbf{H}^{T}(\mathbf{x}^{(0)})\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}^{(0)}) + \mathbf{M}^{-1}\right]^{-1}. \tag{4.18}$$

Combining (4.12), (4.17), and (4.18) at the $(i+1)$-th iteration with an initial point $\mathbf{x}^{(i)} = \mathbf{x}^{(i+1)} - \Delta_{\mathbf{x}}$, the $(i+1)$-th point becomes

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{\Sigma}^{(i)}\{\mathbf{H}^{T}(\mathbf{x}^{(i)})\mathbf{R}^{-1}[\mathbf{z} - \mathbf{h}(\mathbf{x}^{(i)})] - \mathbf{M}^{-1}(\mathbf{x}^{(i)} - \tilde{\mathbf{x}})\}. \tag{4.19}$$

This completes the proof. ■

One main benefit of the state forecasting stage is to provide the initial states to the iterative EKF algorithm in (4.10). Thus, the convergence of the EKF algorithm partly depends on the accuracy of the forecast state vector. A high state forecasting accuracy leads to a faster convergence of the EKF algorithm.

## 4.5 False Data Detection and Identification

The problems of detecting false data injections in the measurement vector and identifying the buses under attack are studied in this section.

Results of the dynamic state estimation will be used in the detection and identification of FDIA. To facilitate the development of the FDIA detection algorithm, let the initial guess $\mathbf{x}^{(i)} = \tilde{\mathbf{x}}$ at time $k+1$ and by performing only one iteration in (4.10), the estimated state vector is approximated as

$$\hat{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_{k+1} + \mathbf{K}_{k+1}\mathbf{v}_{k+1}, \tag{4.20}$$

where

$$\mathbf{v}_{k+1} = \mathbf{z}_{k+1} - \mathbf{h}(\tilde{\mathbf{x}}_{k+1}) \tag{4.21}$$

is the residual vector,

$$\mathbf{K}_{k+1} = \mathbf{\Sigma}_{k+1}\mathbf{H}^T(\tilde{\mathbf{x}}_{k+1})\mathbf{R}_{k+1}^{-1} \tag{4.22}$$

is the gain matrix, and

$$\mathbf{\Sigma}_{k+1} = \left[\mathbf{H}^T(\tilde{\mathbf{x}}_{k+1})\mathbf{R}_{k+1}^{-1}\mathbf{H}(\tilde{\mathbf{x}}_{k+1}) + \mathbf{M}_{k+1}^{-1}\right]^{-1}. \tag{4.23}$$

To facilitate analysis, write the Taylor series expansion of $\mathbf{h}(\mathbf{x}_{k+1})$ around a linearization point $\tilde{\mathbf{x}}_{k+1}$ as

$$\mathbf{h}(\mathbf{x}_{k+1}) = \mathbf{h}(\tilde{\mathbf{x}}_{k+1}) + \mathbf{H}(\tilde{\mathbf{x}}_{k+1})(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}), \tag{4.24}$$

where $\mathbf{H}(\tilde{\mathbf{x}}_{k+1}) = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}|_{\mathbf{x}=\tilde{\mathbf{x}}_{k+1}}$.

The higher order terms of (4.24) are omitted by assuming that the difference $(\mathbf{x}_{k+1}-\tilde{\mathbf{x}}_{k+1})$ is very small.

Combining (4.1), (4.21), and (4.24) gives

$$\mathbf{v}_{k+1} = \mathbf{H}(\tilde{\mathbf{x}}_{k+1})(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}) + \mathbf{e}_{k+1}. \tag{4.25}$$

The covariance matrix, $\mathbf{S}_{k+1}$, of the residual vector, $\mathbf{v}_{k+1}$, can then be calculated as

$$\mathbf{S}_{k+1} = \mathbf{H}(\tilde{\mathbf{x}}_{k+1})\mathbf{M}_{k+1}\mathbf{H}^T(\tilde{\mathbf{x}}_{k+1}) + \mathbf{R}_{k+1}. \tag{4.26}$$

Based on (4.20), the estimated state vector $\hat{\mathbf{x}}_{k+1}$ is a function the residual vector $\mathbf{v}_{k+1}$, the difference between newly received measurements at time $k+1$ and its corresponding predictions $\mathbf{h}(\tilde{\mathbf{x}}_{k+1})$. The newly received measurement vector $\mathbf{z}_{k+1}$ may deviate from its predicated value $\mathbf{h}(\tilde{\mathbf{x}}_{k+1})$. This mismatch between the measured and predicted measurements may be a result of several factors: a sudden change in the system's operating point due to a loss of a large load [16], system faults such as sensor failure or line-to-ground faults, or false data injections in the measurements. The change in the system's operating point is considered as a normal event. However, presence of false data injections is abnormal and can be very harmful to the system. Hence, it is vital to distinguish between state changes due to sudden load change or FDIA, such that false data injections can be detected and removed from the measurements $\mathbf{z}_{k+1}$ before performing state estimations.

We propose a new quickest change detection method by analyzing the statistical properties of the residual vector $\mathbf{v}_{k+1}$. The design criterion of the quickest change detection algorithm is to minimize the worst-case detection delay, subject to the constraint on an upper-bound of the false alarm rate. Specifically, a sudden load change will affect the measurements on all buses based on the physical model of the system. On the other hand,

FDIA will only affect the measurements on a few buses. Thus we propose to distinguish between sudden load change and FDIA by analyzing the statistical correlations of signals from different buses.

### 4.5.1 Formulation of the Hypothesis Test

Define the null hypothesis $\mathcal{H}_0$, which corresponds to the measurements without false data at time $k+1$, and the alternative hypothesis $\mathcal{H}_1$, which corresponds to the measurements with false data at time $k+1$, as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{z}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{e}_{k+1}, \\ \mathcal{H}_1 &: \mathbf{z}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{e}_{k+1} + \mathbf{a}, \end{aligned} \tag{4.27}$$

where $\mathbf{a}$ is a vector of false data injected in the measurements.

From (4.21), (4.25), and (4.27), the hypothesis test on the residual vector $\mathbf{v}_{k+1}$ can be written as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{v}_{k+1} = \mathbf{H}(\tilde{\mathbf{x}}_{k+1})(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}) + \mathbf{e}_{k+1}, \\ \mathcal{H}_1 &: \mathbf{v}_{k+1} = \mathbf{H}(\tilde{\mathbf{x}}_{k+1})(\mathbf{x}_{k+1} - \tilde{\mathbf{x}}_{k+1}) + \mathbf{e}_{k+1} + \mathbf{a}. \end{aligned} \tag{4.28}$$

The residual vector $\mathbf{v}_{k+1}$ under the null hypothesis $\mathcal{H}_0$ is generally assumed to be a zero mean Gaussian vector [14] and [19]. With the dynamic state estimator presented in this chapter, the covariance matrix $\mathbf{S}_{k+1}$ of the residual vector is given in (4.26). Assuming that the attack vector $\mathbf{a}$ is a deterministic vector, under the alternate hypothesis $\mathcal{H}_1$, $\mathbf{v}_{k+1}$ is Gaussian with mean $\mathbf{a}$ and covariance matrix $\mathbf{S}_{k+1}$.

As in (4.26), the elements in $\mathbf{z}_{k+1}$ are correlated based on the physical model of the power grid. To simplify the analysis, we propose to perform a whitening transformation on $\mathbf{v}_{k+1}$.

The covariance matrix of the residual vector can be decomposed as $\mathbf{S}_{k+1} = \mathbf{U}_{k+1}^T \mathbf{D}_{k+1} \mathbf{U}_{k+1}$, where $\mathbf{D}_{k+1}$ is a diagonal matrix with the eigenvalues of $\mathbf{S}_{k+1}$ on its main diagonal and $\mathbf{U}_{k+1}$ is the corresponding orthonormal eigenvector matrix at time instant $k+1$. The whitening transformation of the residual vector $\mathbf{v}_{k+1}$ is $\bar{\mathbf{v}}_{k+1} = \mathbf{W}_{k+1}\mathbf{v}_{k+1}$, where the whitening matrix $\mathbf{W}_{k+1} = \mathbf{D}_{k+1}^{-\frac{1}{2}} \mathbf{U}_{k+1}$. With the whitening operator, it can be easily shown that the covariance matrix of $\bar{\mathbf{v}}_{k+1}$ is $\mathbf{I}_m$, which is an $m \times m$ identity matrix.

Following the Gaussian distribution of $\mathbf{v}_{k+1}$ given in (4.28), the hypothesis test on $\bar{\mathbf{v}}_{k+1}$ is

$$\begin{aligned} \mathcal{H}_0 : \bar{\mathbf{v}}_{k+1} = \mathbf{W}_{k+1}\mathbf{v}_{k+1} \quad &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}_m), \\ \mathcal{H}_1 : \bar{\mathbf{v}}_{k+1} = \mathbf{W}_{k+1}\mathbf{v}_{k+1} \quad &\sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{I}_m), \end{aligned} \tag{4.29}$$

where $\boldsymbol{\mu} = \mathbf{W}_{k+1}\mathbf{a}$.

### 4.5.2 Proposed False Data Detector

A QCD-based false data detection method is proposed in this section to detect cyber-attacks. We assume that the false data is injected at a random time $\tau$, and the attack was detected at time $\hat{\tau}$. Based on the design criteria of quickest change detection, the problem can be formulated as

$$\begin{aligned} \text{(P1)} \qquad \text{minimize} \quad & \text{WDD} = \sup_k \mathbb{E}_k[(\hat{\tau} - k)^+] \\ \text{subject to} \quad & \text{FAR} = \frac{1}{\mathbb{E}_\infty[\hat{\tau}]} \leq \zeta. \end{aligned}$$

where WDD is the worst-case detection delay, FAR is the false alarm rate, $(a)^+ = a$ if $a \geq 0$ and 0 otherwise, $\mathbb{E}_k$ is the expectation assuming the attack becomes active at $\tau = k$, and $\mathbb{E}_\infty$ denotes the expectation when there is no attack. The solution of the problem is a quickest

detection algorithm in that it aims at minimizing the worst-case detection delay, subject to an upper bound of the false alarm rate.

The above problem can be solved by using the well-known CUSUM algorithm [27]

$$\hat{\tau} = \inf\{k \geq 1 | C_k \geq A\}, \tag{4.30}$$

where $A$ is a threshold obtained by the FAR upper bound $\zeta$,

$$C_{k+1} = \max(0, C_k + L_k), \tag{4.31}$$

and $L_k = \log \frac{f_1(\bar{\mathbf{v}}_k)}{f_0(\bar{\mathbf{v}}_k)}$ is the log-likelihood ratio (LLR), with $f_1(\bar{\mathbf{v}}_k)$ and $f_0(\bar{\mathbf{v}}_k)$ being the probability density functions (pdfs) associated with hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively. The CUSUM algorithm is the asymptotically optimum solution to (P1) because it can asymptotically minimize the WDD when the FAR goes to 0 [28, 29].

Under the assumption that $\bar{\mathbf{v}}_k$ is Gaussian distributed, the LLR can be calculated as

$$L_k = \mathbf{a}^T \mathbf{W}_k^T \bar{\mathbf{v}}_k - \frac{1}{2} \mathbf{a}^T \mathbf{W}_k^T \mathbf{W}_k \mathbf{a}. \tag{4.32}$$

The calculation of the LLR $L_k$ requires the knowledge of the attack vector $\mathbf{a}$, which is unknown at the detector. Thus we cannot directly apply the CUSUM algorithm. In order to resolve the unknown parameters, the detection method in [29] utilizes the generalized likelihood ratio test (GLRT) approach by replacing the unknown parameter with the maximum likelihood estimation (MLE) as

$$\hat{\tau} = \inf\{k \geq 1 | \max_{1 \leq t \leq k} \sup_{\mathbf{a}} \sum\nolimits_{i=t}^{k} L_i \geq A\}. \tag{4.33}$$

This approach is proven to be asymptotically optimal in terms of minimum detection delay [34]. However, the test statistic cannot be computed recursively as the CUSUM test,

because GLRT needs to compute every unknown element of $\mathbf{a}$ for each observation at sampling time $1 \leq t \leq k$. In other words, GLRT needs to store the observations and perform MLE of $\mathbf{a}$ at every sampling instant. As a result, this approach has very high complexity, and it might not be feasible for real-time FDIA detection in power grids.

A low-complexity adaptive-CUSUM algorithm is proposed in [8] based on Rao test [31], which is asymptotically equivalent to the GLRT test [35]. The Rao test statistic can be computed by taking the derivative of $L_k$ with respect to the unknown parameter $\mathbf{a}$ evaluated around the region of interests. In our case, the region of interest is considered to be around zero because the hypothesis $\mathcal{H}_0$ has zero mean. According to (4.29), the statistic [31] of the Rao test for detection at time $k$ can be written as follows:

$$Y(\bar{\mathbf{v}}_k) = \frac{\partial L_k}{\partial \mathbf{a}}^T \Big|_{\mathbf{a}=0} \left[\mathcal{I}^{-1}(\mathbf{a})\big|_{\mathbf{a}=0}\right] \frac{\partial L_k}{\partial \mathbf{a}}\Big|_{\mathbf{a}=0} = \bar{\mathbf{v}}_k^T \bar{\mathbf{v}}_k, \tag{4.34}$$

where $\mathcal{I}(\mathbf{a})$ is the Fisher information matrix [23]. The proof for (4.34) is given below.

*Proof:* Combining the definition of LLR in (4.32) with the hypotheses in (4.29), we obtain

$$\frac{\partial L_k}{\partial \mathbf{a}} = \mathbf{W}_k^T \bar{\mathbf{v}}_k - \mathbf{W}_k^T \mathbf{W}_k \mathbf{a}. \tag{4.35}$$

Next, substituting the value $\mathbf{a} = 0$ yields in

$$\frac{\partial L_k}{\partial \mathbf{a}}\Big|_{\mathbf{a}=0} = \mathbf{W}_k^T \bar{\mathbf{v}}_k. \tag{4.36}$$

Using the definition of Fisher information matrix [23], we get

$$\mathcal{I}(\mathbf{a}) = -\mathbb{E}\left[\frac{\partial}{\partial \mathbf{a}}\left(\frac{\partial L_k}{\partial \mathbf{a}}\right)\right] = \mathbf{W}_k^T \mathbf{W}_k. \tag{4.37}$$

Combining (4.36) and (4.37) results in the Rao test statistic $Y(\bar{\mathbf{v}}_k) = \bar{\mathbf{v}}_k^T \bar{\mathbf{v}}_k$. This completes the proof. ■

Under the null hypothesis $\mathcal{H}_0$, the Rao test statistic follows Chi-square distribution, that is, $Y(\bar{\mathbf{v}}_k) \sim \chi^2_m$, where $m$ is the degree-of-freedom corresponding to the dimension of the measurement vector. If we directly replace the LLR $L_k$ in (4.31) with the Rao test statistic $Y(\bar{\mathbf{v}}_k)$ in the CUSUM test defined in (4.30), the CUSUM test statistic $C_k$ will increase monotonically under both the null and alternative hypothesis because $Y(\bar{\mathbf{v}}_k) \geq 0$. This is undesirable for CUSUM because it is a threshold test. To address this issue, we introduce a normalized version of the test statistic in (4.34) with respect to the mean and standard deviation of $Y(\bar{\mathbf{v}}_k)$ under $\mathcal{H}_0$, which are, $m$ and $\sqrt{2m}$, respectively. Based on the normalized test statistic, we propose a new detection rule as follows.

*Definition 4.1*: (Normalized Rao-CUSUM Detector) Given a whitened residual vector $\bar{\mathbf{v}}_{k+1} = \mathbf{W}_{k+1}\mathbf{v}_{k+1}$ at time $k+1$ , an FDIA is detected at time $\hat{\tau}$ with

$$\hat{\tau} = \inf\{k \geq 1 | T_k \geq A\}, \tag{4.38}$$

where

$$T_{k+1} = \max\left(0, T_k + \frac{Y(\bar{\mathbf{v}}_{k+1}) - m}{\sqrt{2m}}\right), \tag{4.39}$$

with $T_0 = 0$. The threshold $A$ is determined by the FAR upper bound $\zeta$.

The normalized Rao-CUSUM detector is developed by modifying the asymptotically optimum GLRT-CUSUM detector to balance the tradeoff between complexity and performance. The proposed normalized Rao-CUSUM algorithm might have a bit higher WDD than the GLRT-CUSUM algorithm, but offers much lower complexity.

It should be noted that the above test can distinguish between sudden load change from FDIA because the formulation of the null hypothesis $\mathcal{H}_0$ includes sudden load change as a special case. In case of a sudden load change, the system dynamics still follow the physical

model of the power grid. As a result, the residual vector can still be modeled as zero-mean Gaussian distributed with covariance matrix $\mathbf{S}_{k+1}$. Yet this is no longer true when there is false data injected into the power grid, which is modeled as the alternative hypothesis $\mathcal{H}_1$. Since the test in Definition 1 is designed to distinguish between the null and alternative hypothesis by minimizing the worst-case detection delay, it is able to distinguish between load change and FDIA. On the other hand, system faults, such as sensor failures or line faults, will also cause the measurements to deviate from those predicted by the physical model of the system. In that case, the proposed algorithm will be able to detect the presence of sensor failures or system faults. However, it will not be able to differentiate FDIA from sensor failures or system faults. Thus the algorithm will treat FDIA, sensor failure, or other system faults in a similar manner.

In case false data are detected, we can identify the buses under attack by using the power of the residuals at different buses. That is, if the residual power or amplitude on a given bus is above a certain threshold, then it is considered that the corresponding bus is under attack. Similar to [14], the amplitude test can be expressed as

$$|\mathbf{v}_{k+1}(i)| > \gamma\sigma_{S_i}, \tag{4.40}$$

where $|\mathbf{v}_{k+1}(i)|$ is the absolute value of the $i$-th element of $\mathbf{v}_{k+1}$, $\sigma_{S_i}$ is the standard deviation of the $i$-th element of $\mathbf{v}_{k+1}$, and $\gamma$ defines the limit of confidence. If a bus is detected as under attack, we replace the estimated states with the predicted states to ensure the normal operations of the power grid.

The normalized Rao-CUSUM detector proposed in Definition 1 is a simple threshold test, and the test statistic $T_k$ can be recursively calculated based on (4.39). As a result, the

proposed test has low complexity and can be easily implemented. The implementation of the detector in (4.38) requires a threshold $A$, which in turn depends on the FAR upper bound $\zeta$. In the next section, we will provide a theoretical guideline for choosing the threshold value $A$ in terms of FAR with the help of a Markov-chain-based analytical model.

### 4.5.3 Computation Complexity Analysis

In this subsection, we study the effects of the size of system on the computation complexity of the proposed algorithm. The size of the system can be defined by two parameters:

(1) The dimension of state vector: $n = 2N - 1$, where $N$ is the number of buses,

(2) The dimension of measurement vector: $m$, which depends on the number of buses and lines.

It is easily observed that $m > n$. To determine the effect of size of the system on the performance, we present the complexity analysis of the proposed algorithm in a single sampling instant $k$ with respect to these two parameters separately.

The proposed algorithm has two stages: 1) false data detection and 2) dynamic state estimation. With respect to $m$, computation in stage 1 is dominated by the eigenvalue decomposition process which has a cubic complexity $\mathcal{O}(m^3)$, and stage 2 is dominated by the matrix inversion of $\mathbf{R}_k$ which also has a cubic complexity $\mathcal{O}(m^3)$. So, total complexity of the proposed algorithm scales cubically with $m$ as $\mathcal{O}(m^3)$.

With respect to $n$, stage 2 computation is dominated by the computations of $\mathbf{M}_k$ in (4.8) and $\mathbf{\Sigma}_k$ in (4.11), both of which have a cubic complexity $\mathcal{O}(n^3)$. Comparatively, complexity of stage 1 scales quadratically with $n$. So, total complexity of the proposed algorithm scales cubically with $n$ as $\mathcal{O}(n^3)$.

## 4.6 Markov-chain-based Analytical Model

In this section, we present a Markov-chain-based model to analyze the proposed false data detector. The Markov-chain-based model provides theoretical guidelines on the choice of the detection threshold in (4.38) based on the FAR upper bound. For a given FAR upper bound, we can obtain the optimum detection threshold by using offline Monte-Carlo simulations. Once the optimum threshold is obtained offline, the online normalized Rao-CUSUM detector can then be performed to detect FDIA or system faults in real time.

To facilitate analysis, $\mathbb{R}^+ \cup\ 0$ is discretized into a finite set of intervals representing the states $\{U_0, U_1, \cdots, U_M\}$ such as

$$U_0 = 0, \quad U_1 = (0, \Delta], \quad U_2 = (\Delta, 2\Delta],$$
$$U_3 = (2\Delta, 3\Delta], \quad \cdots\cdots, \quad U_M = (A, +\infty),$$

where $\Delta = \frac{A}{M-1}$ and $M$ represents the total number of transitions from 0 to the state that has the value greater than the threshold $A$.

It can be easily observed from (4.39) that the sequence exhibits the property of a first-order Markov chain, where the future state $T_{k+1}$ at time index $k+1$ depends only on the current state $T_k$, but not on past states [36].

The transition probabilities of the Markov chain under $\mathcal{H}_0$ for the proposed algorithm from state $U_i$ at $k$ to state $U_j$ at $k+1$ can be described as

$$P_{ij} = P(T_{k+1} \in U_j | T_k \in U_i). \tag{4.41}$$

The transition probability $P_{ij}$ can be computed numerically using Monte-Carlo simulations according to the distribution of $\bar{\mathbf{v}}$ under the null hypothesis $\mathcal{H}_0$. The values of the transition

probabilities are uniquely determined by the threshold $A$ and the number of discretization levels $M$. Since the calculations of the transition probabilities are performed offline, we can achieve arbitrary precision of the transition probability by increasing the number of Mote-Carlo trials without affecting the complexity of the online portion of the algorithm. Therefore, we can establish a very accurate numerical relationship between $A$ and the transition probabilities.

Define the transition probability matrix (TPM) $\mathbf{P}$ as an $(M+1)\times(M+1)$ matrix with the $(i,j)$-th element being $P_{i-1,j-1}$. It is clear that $\mathbf{P}$ is a Markov matrix, that is, all elements of $\mathbf{P}$ are non-negative and the sum of each row vector is 1. The steady-state probability $\pi_j$ of each state $U_j$ can be determined by

$$\pi_j = \sum\nolimits_{i=0}^{M} P_{ij}\pi_i,\ \forall\ j \in \{0,\cdots,M\}, \tag{4.42}$$

$$\sum\nolimits_{j=0}^{M} \pi_j = 1. \tag{4.43}$$

The transition probability of the Markov chain can be written in a matrix format as

$$\mathbf{P}^T\boldsymbol{\pi} = \boldsymbol{\pi} \tag{4.44}$$

where $\boldsymbol{\pi} = [\pi_0, \pi_1, \cdots, \pi_M]^T$. The steady-state probability vector $\boldsymbol{\pi}$ can then be obtained by finding the eigenvector corresponding to the eigenvalue 1 of the TPM $\mathbf{P}$. Since $\mathbf{P}$ is a Markov matrix, it always has an eigenvalue of 1.

The steady-state probability can be used to calculate the FAR, which can be equivalently evaluated as the probability that $T_k$ crosses the threshold $A$ when there is no attack in the network. As in [36], the FAR can be equivalently calculated as the steady-state probability

$\pi_M$, that is, the probability that $T_k$ stays at state $U_M$ under the null hypothesis

$$\text{FAR} = \pi_M. \tag{4.45}$$

Since $\pi_M$ is determined by the eigenvector of $\mathbf{P}$, which in turn depends on the choice of threshold $A$, there is an optimum threshold value for a given FAR. Enabled by the Markov-chain model, we can numerically obtain a very accurate estimate of the optimum threshold based on the FAR.

## 4.7 Simulation Results

In this section, we present numerical simulations results to illustrate the performance of the proposed algorithm. The first subsection demonstrates the performance in terms of FAR and WDD using simulated data. The second subsection presents numerical results based on simulations of the 13-bus system using MATLAB Power System Toolbox (PST v3.0) [37,38].

### 4.7.1 WDD v.s. FAR

Figure 4.1 shows the tradeoff between WDD and FAR of the proposed algorithm and the Rao-CUSUM test [8]. In the simulation, the data are generated by following the model in (4.29) with $m = 55$ and $\boldsymbol{\mu} = [1, 1, 0, \cdots, 0]$. The false data injection time $\tau$ follows discrete uniform distribution between $[1, 100]$. Every point on the curves is obtained by averaging over $10,000$ Monte-Carlo trials. For a given FAR, the corresponding threshold $A$ is chosen by following the Markov-chain analysis in Section 4.6. As expected, the WDD is a decreasing function of the FAR. The proposed detection algorithm outperforms the Rao-CUSUM test used in [8]. At FAR $= 10^{-2}$, the WDD of the proposed algorithm and the Rao-CUSUM test is 42 and 95 samples, respectively.

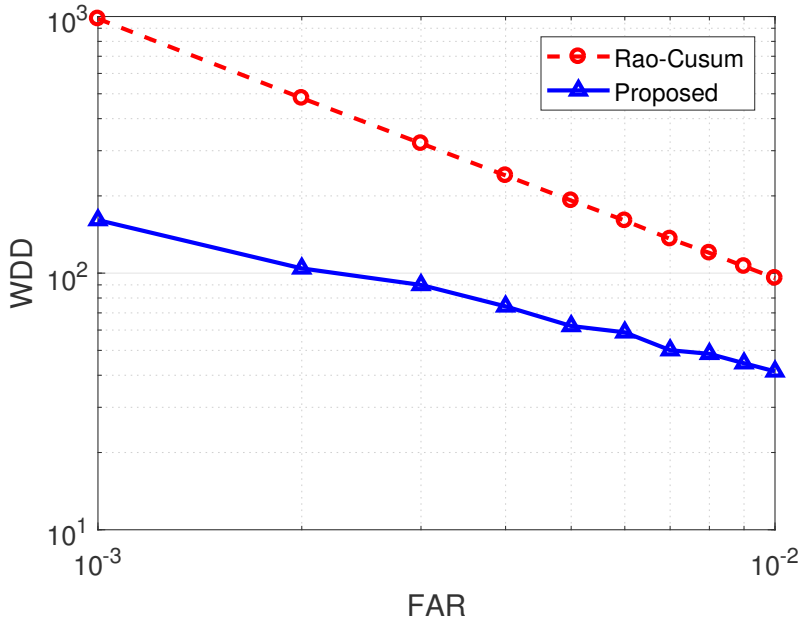


Figure 4.1: Performance analysis of the proposed algorithm in comparison with Rao-CUSUM test.

### 4.7.2 FDIA Detection in Power Systems

In this section, we present the simulation results performed on a 13-bus system with two areas as shown in Figure 4.2 [38]. Bus 1 is used as the reference bus. The measurement vector consists of $m = 55$ components: the voltage magnitude of bus 1, the active and reactive power injections at all 13 buses, the active and reactive power flows at all 14 lines. The state vector consists of $n = 25$ components: the voltage magnitudes at all 13 buses and the phase angles at the 12 non-reference buses.

Using MATLAB Power System Toolbox (PST v3.0), the system dynamics is simulated by increasing the active load at bus 4 by 0.5 per unit (p.u.) and the resulting measurement and state vectors are considered as the true values of the system. The noisy measurement vector $\mathbf{e}_k$ in (4.1) is obtained by adding a zero mean Gaussian noise with a diagonal covariance matrix $\mathbf{R}_k$ to each of the true measurements. The noise variances, which are the diagonal

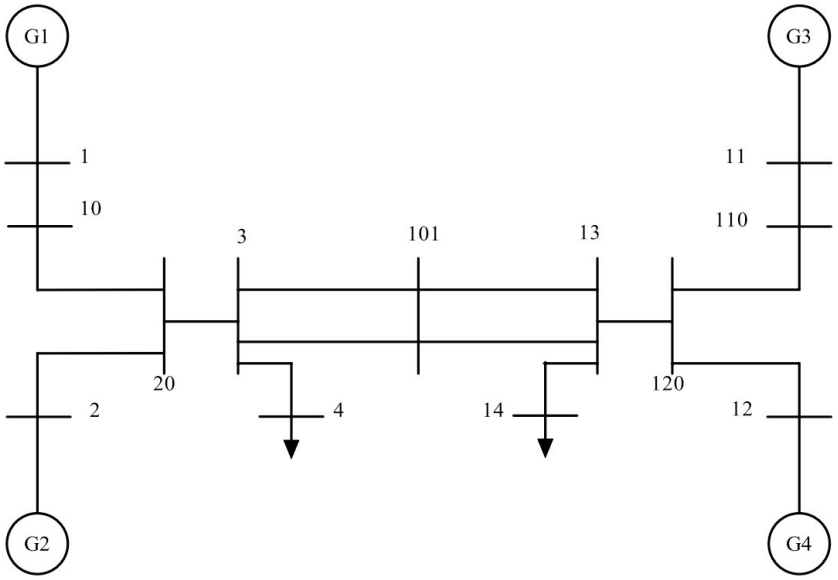


Figure 4.2: Single Line Diagram Two Area System.

elements of $\mathbf{R}_k$, are $10^{-5}$ for the voltage magnitude of reference bus and $10^{-6}$ for the active and reactive power measurements. The matrix $\mathbf{Q}_k = 10^{-6}\mathbf{I}_n$ is kept constant at every sampling time $k$. The parameters $\mathbf{F}_k$ and $\mathbf{G}_k$ are obtained according to the Holt's exponential smoothing method with $\alpha = 0.95$ and $\beta = 0.001$ [14].

The sampling rate in our simulation is set as $\Delta t = 0.01$ seconds. Thus the $k$-th time index corresponds to a time value of $t = k\Delta t$ seconds. In order to evaluate the performance of the proposed detector, two scenarios are simulated: false data and sudden load change conditions. The false data condition is simulated by injecting errors of $-1.5$ and 1 p.u. into the active power measurements at buses 13 and 14, respectively, during a time period $0.25 \leq t < 0.6$ seconds unless specified otherwise. The sudden load change condition is simulated by cutting the active power injection of bus 4 by 1 p.u at $t = 0.6$ seconds. In each of the following figures, every point on the curves is obtained by averaging over $1,000$ Monte-Carlo trials.

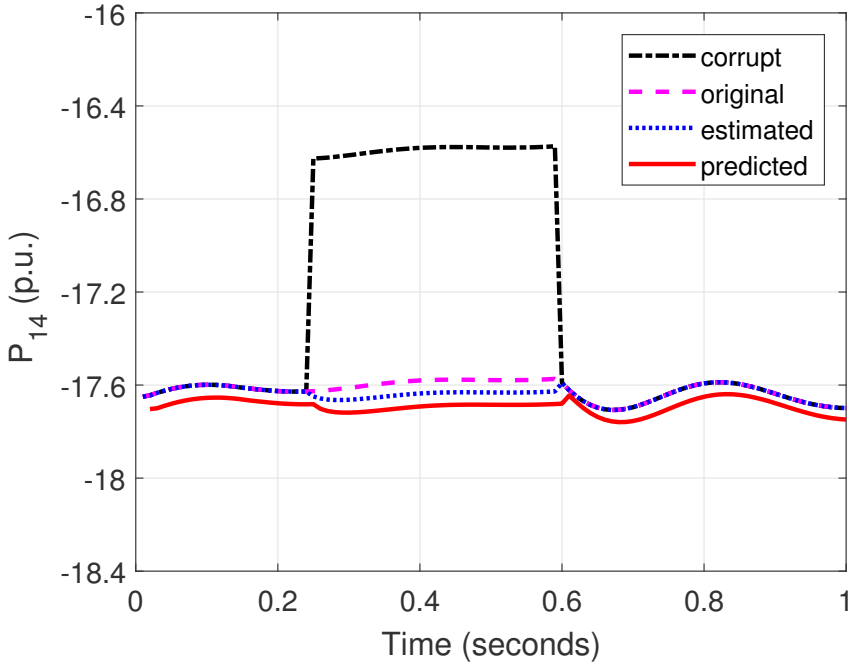


Figure 4.3: The real power at bus 14 vs time $t$ with false data at $0.25 \leq t < 0.6$, load change at $t = 0.6$, and the proposed detector in (4.39).

Figure 4.3 shows the active power at bus 14 with false data injected into the active power measurements at buses 13 and 14. In addition, the active power injection of bus 4 is cut by 1 p.u. at $t = 0.6$ seconds to simulate sudden load change. The threshold of the proposed detector is set at $A = 200$, which corresponds to FAR of $2.5 \times 10^{-5}$ according to the Markov-chain analysis. Once an FDIA is detected, the residual amplitudes are compared to a threshold as in (4.40) to identify the buses under attack, with $\gamma = 3.5$ used in this chapter. The measurements at the buses under attack are then replaced with their predicted values. When there is no attack, the power calculated from the estimated states is almost identical to its actual value. When false data is injected between $0.25 \leq t < 0.6$ seconds, the proposed detector successfully detects the presence of FDIA and replaces the corrupted measurement with the predicted values. In this case, the power calculated by using state estimation is slightly different from its true value, with a difference less than 0.33%. When

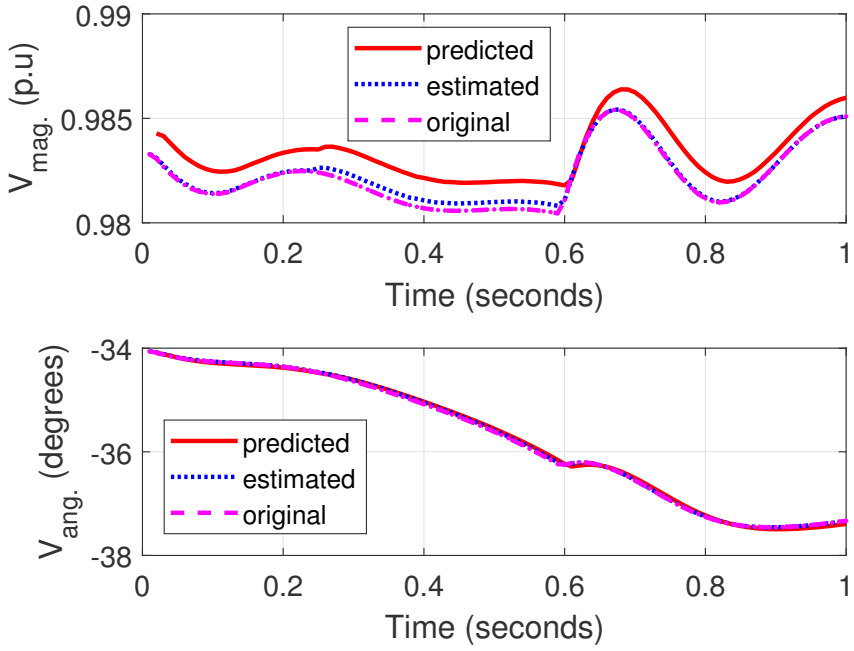


Figure 4.4: The voltage magnitude (top) and phase angle (bottom) at bus 13 vs time $t$ with false data at $0.25 \leq t < 0.6$, load change at $t = 0.6$, and the proposed detector in (4.39).

there is a sudden change at $t = 0.6$ seconds, the detector correctly recognizes it as a normal operating condition and achieves correct state estimates. In addition, there is a one sample lag between the predicted value and the actual value.

Figure 4.4 shows the voltage magnitude (top) and phase angle (bottom) at bus 13 under the same configuration of Figure 4.3. The voltage amplitude and phase are estimated with high accuracy despite the presence of FDIA, mainly because the false data are correctly identified and replaced with predicted values. In addition, the state estimator correctly adapts to the dynamic change at $t = 0.6$ seconds.

Figure 4.5 shows the performance of an existing residual-based detector that was proposed in [14], under the same configuration as in Figure 4.3. Since the false data are injected into the correlated measurements of adjacent buses 13 and 14, the detector in [14] is unable to distinguish the false data from the sudden changes in the system. The FDIA is erroneously

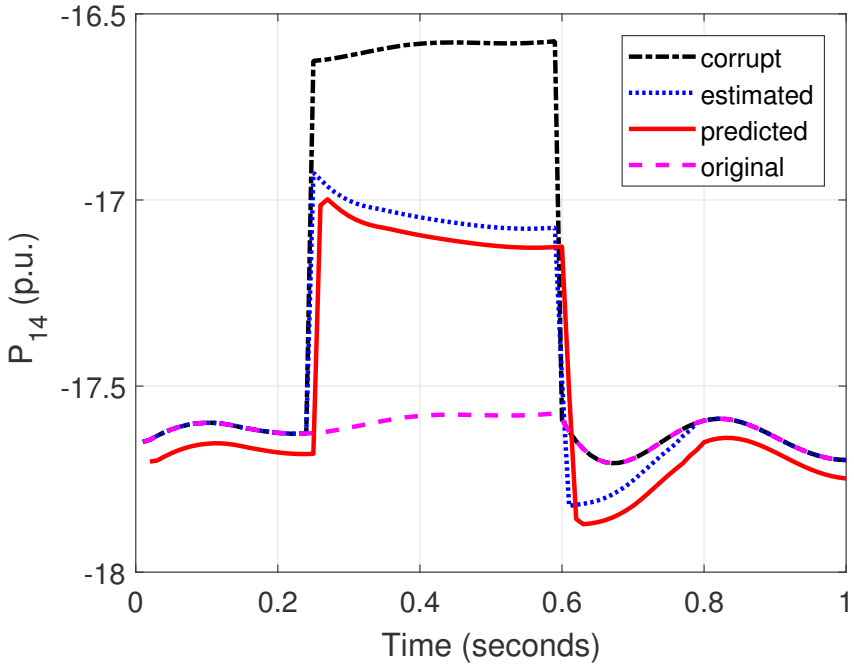


Figure 4.5: The real power at bus 14 vs time $t$ with false data at $0.25 \leq t < 0.6$ and the conventional detector.

detected as a sudden change. As a result, during the FDIA, the power calculated from the estimated and predicted states deviate significantly from its actual value. The performance of the estimator yields in an estimation error of as high as 0.7 p.u at $t = 0.25$ seconds.

To further illustrate the ability of the proposed detector to distinguish between FDIA and sudden change, Figure 4.6 shows the active power measurement at bus 4. The active load at bus 4 is increased by 0.5 p.u. at $t = 0.6$ seconds. The load change caused a gradual change of the active power. Since the load change affects power measurements on all buses based on the physical model of the power grid, the proposed detector successfully recognizes it as a load change instead of FDIA. Thus the dynamic state estimator can accurately track and estimate the state change caused by the load change. The power calculated from the estimated states is almost identical to its original value. Again a one sample lag is observed between the predicted value and actual value. Similarly, in Figure 4.7, the voltage magnitude

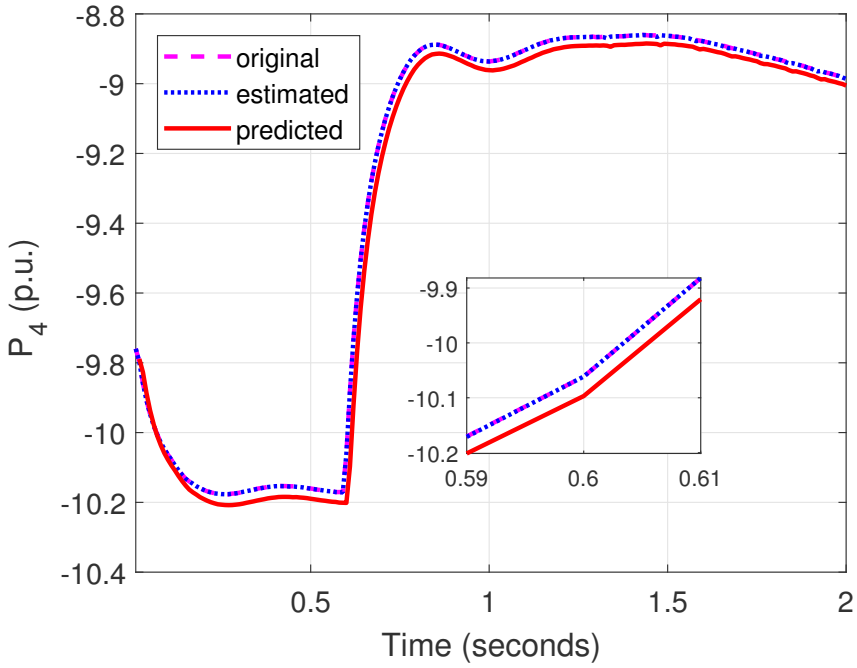


Figure 4.6: The real power at bus 4 vs time $t$ with false data at $0.25 \leq t < 0.6$, load change at $t = 0.6$, and the proposed detector in (4.39).

(top) and phase angle (bottom) at bus 4 are estimated with high accuracy.

The proposed algorithm assumes that the topology of the system remains unchanged. However, in the event of a system fault, the topology of the system might change. As a result, the proposed algorithm will detect the deviation of system's behavior due to system fault in a similar manner as FDIA detection. To further illustrate the performance of proposed algorithm under system faults, Figure 4.8 shows the active power at bus 14 under the influence of a single line-to-ground fault, which is applied to the line connecting bus 3 and bus 101 at $t = 0.2$ seconds. The fault is cleared at bus 3 at $t = 0.35$ seconds and at bus 101 at $t = 0.4$ seconds. It can be observed that throughout the duration of the line-to-ground fault, the proposed algorithm detects the fault and replace the measurement value by using the predicted values. Thus the algorithm can detect the presence of fault, but it cannot differentiate fault from cyber-attacks.

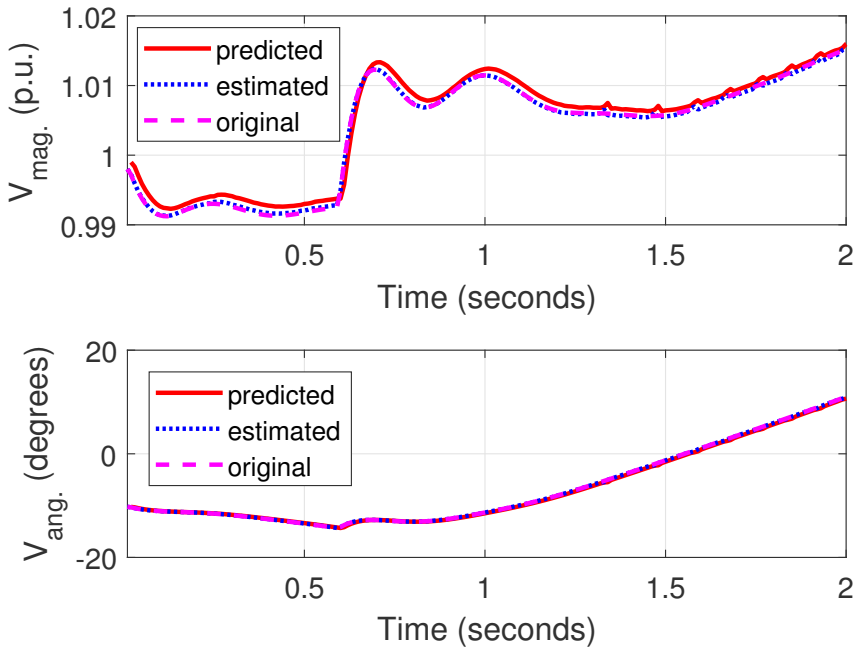


Figure 4.7: The voltage magnitude (top) and phase angle (bottom) at bus 4 vs time $t$ with false data at $0.25 \leq t < 0.6$, load change at $t = 0.6$, and the proposed detector in (4.39).

## 4.8 Conclusion

A quickest intrusion detection algorithm has been developed for the detection and removal of false data injected into smart grids. The algorithm was developed to minimize the worst-case detection delay subject to an upper bound of false alarm rate. To distinguish between FDIA and sudden system change, a time-varying dynamic model was used to represent the dynamic state transitions. A dynamic state estimator was then developed to estimate and track the time-varying and non-stationary state transitions. Based on the statistical properties of the state estimation results, a new normalized Rao-CUSUM detector was developed to minimize the detection delay of FDIA while separating FDIA from sudden system changes. Unlike existing algorithms that rely on measurement correlation to discriminate false data from sudden system changes, the proposed algorithm can detect any false data including those

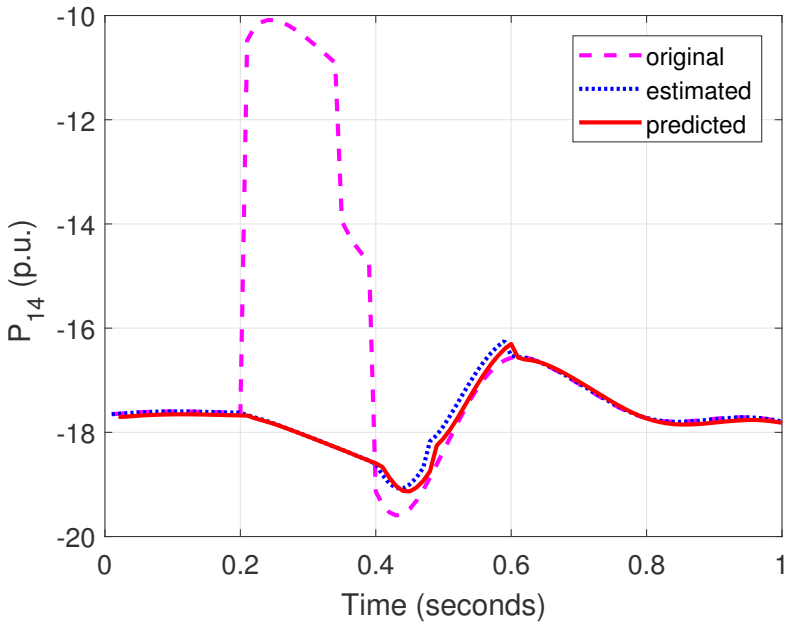


Figure 4.8: The real power at bus 14 with single line-to-ground fault at the line connecting bus 3 and bus 101 during $0.2 \leq t \leq 0.4$.

injected into correlated measurements. Simulation results have shown that the proposed algorithm can accurately and timely detect and remove FDIA. In addition, the algorithm can also detect system faults such as sensor failures or line outages. However, it cannot differentiate system faults from FDIA. The algorithm can be used to harden IEDs or SCADA systems to improve the security and resilience of smart grids.

## References

[1] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, vol. 7, no. 3, pp. 75–77, May 2009.

[2] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 106–115, Sep. 2012.

[3] D. Wei, Y. Lu, M. Jafari, P. M. Skare, and K. Rohde, "Protecting smart grid automation systems against cyberattacks," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 782–795, Dec.

2011.

[4] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security & Privacy*, vol. 9, no. 3, pp. 49–51, May 2011.

[5] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.

[6] E. Pallotti and F. Mangiatordi, "Smart grid cyber security requirements," in *IEEE International Conference on Environment and Electrical Engineering*, Rome, Italy, May 2011.

[7] A. J. McBride and A. R. McGee, "Assessing smart grid security," *Bell Labs Technical Journal*, vol. 17, no. 3, pp. 87–103, Dec. 2012.

[8] Y. Huang, J. Tang, Y. Cheng, H. Li, K. A. Campbell, and Z. Han, "Real-time detection of false data injection in smart grid networks: an adaptive cusum method and analysis," *IEEE Systems Journal*, vol. 10, no. 2, pp. 532–543, Jun. 2016.

[9] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 13:1–13:33, May 2011.

[10] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. on Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[11] Z.-H. Yu and C. Wen-Long, "Blind false data injection attack using pca approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.

[12] J. Liang, O. Kosut, and L. Sankar, "Cyber attacks on ac state estimation: Unobservability and physical consequences," in *IEEE PES General Meeting— Conference & Exposition*. National Harbor, MD, USA: IEEE, Jul. 2014.

[13] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks against nonlinear state estimation in smart power grids," in *IEEE PES General Meeting*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5.

[14] A. L. Da Silva, M. Do Coutto Filho, and J. De Queiroz, "State forecasting in electric power systems," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 130, no. 5, Sep. 1983, pp. 237–244.

[15] K.-R. Shih and S.-J. Huang, "Application of a robust algorithm for dynamic state estimation of a power system," *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 141–147, Feb. 2002.

[16] M. B. Do Coutto Filho and J. C. S. de Souza, "Forecasting-aided state estimation—part i: Panorama," *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1667–1677, Nov. 2009.

[17] H. Li and W. Li, "Estimation and forecasting of dynamic state estimation in power systems," in *IEEE International Conference on Sustainable Power Generation and Supply*, Nanjing, China, Apr. 2009.

[18] A. Jain and N. Shivakumar, "Power system tracking and dynamic state estimation," in *IEEE/PES Power Systems Conference and Exposition*, Seattle, WA, USA, Mar. 2009.

[19] K. Nishiya, J. Hasegawa, and T. Koike, "Dynamic state estimation including anomaly detection and identification for power systems," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 129, no. 5, Sep. 1982, pp. 192–198.

[20] Q. Yang, L. Chang, and W. Yu, "On false data injection attacks against kalman filtering in power system dynamic state estimation," *Security and Communication Networks*, vol. 9, no. 9, pp. 833–849, Jun. 2016.

[21] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting false data injection attacks in ac state estimation," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2476–2483, Sep. 2015.

[22] J. James, Y. Hou, and V. O. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, Jul. 2018.

[23] S. M. Kay, *Fundamentals of statistical signal processing*. Prentice Hall PTR, 1993.

[24] I. Akingeneye and J. Wu, "Low latency detection of sparse false data injections in smart grids," *IEEE Access*, vol. 6, pp. 58 564–58 573, Oct. 2018.

[25] H. V. Poor and O. Hadjiliadis, *Quickest detection*. Cambridge University Press Cambridge, 2009, vol. 40.

[26] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. Chapman and Hall/CRC, 2014.

[27] G. V. Moustakides *et al.*, "Optimal stopping times for detecting changes in distributions," *The Annals of Statistics*, vol. 14, no. 4, pp. 1379–1387, Dec. 1986.

[28] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.

[29] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[30] S. Nath and J. Wu, "Bayesian quickest change point detection with multiple candidates of post-change models," in *IEEE Global Conference on Signal and Information Processing*, Anaheim, CA, USA, Nov. 2018.

[31] A. De Maio, "Rao test for adaptive detection in gaussian interference with unknown covariance matrix," *IEEE transactions on signal processing*, vol. 55, no. 7, pp. 3577–3584, Jul. 2007.

[32] X. Li, A. Scaglione, and T.-H. Chang, "Optimal sensor placement for hybrid state estimation in smart grid," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, May 2013.

[33] S. Akram and Q. ul Ann, "Newton raphson method," *International Journal of Scientific & Engineering Research*, vol. 6, no. 7, pp. 1748–1752, Jul. 2015.

[34] T. L. Lai and J. Z. Shan, "Efficient recursive algorithms for detection of abrupt changes in signals and control systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 5, pp. 952–966, 1999.

[35] A. De Maio and S. Iommelli, "Coincidence of the rao test, wald test, and glrt in partially homogeneous environment," *IEEE Signal Processing Letters*, vol. 15, pp. 385–388, Apr. 2008.

[36] D. Gamerman and H. F. Lopes, *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman and Hall/CRC, 2006.

[37] J. H. Chow and K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE transactions on Power Systems*, vol. 7, no. 4, pp. 1559–1564, Nov. 1992.

[38] G. Rogers, *Power system oscillations*. Springer Science & Business Media, 2012.

# Chapter 5

## Low Latency Bearing Fault Detection of Direct-drive Wind Turbines Using Stator Current

Samrat Nath, Jingxian Wu, Yue Zhao, Wei Qiao

### 5.1 Abstract

Low latency change detection aims to minimize the detection delay of an abrupt change in probability distributions of a random process, subject to certain performance constraints such as the probability of false alarm (PFA). In this chapter, we study the low latency detection of bearing faults of direct-drive wind turbines (WT), by analyzing the statistical behaviors of stator currents generated by the WT in real-time. It is discovered that the presence of fault will affect the statistical distribution of WT stator current amplitude at certain frequencies. Since the signature of a fault can appear in one of the multiple possible frequencies, we need to monitor the signals on multiple frequencies simultaneously, and each possible frequency is denoted as a candidate. Based on the unique properties of WT bearing faults, we propose a new multi-candidate low latency detection algorithm that can combine the statistics of signals from multiple candidate frequencies. The new algorithm does not require a separate training phase, and it can be directly applied to the stator current data and perform online detection of various possible bearing faults. The theoretical performance of the proposed algorithm is analytically identified in the form of upper bounds of the PFA and average detection delay (ADD). The algorithm allows flexible parametric adjustment of the tradeoff between PFA and ADD.

## 5.2 Introduction

Wind is one of the top sources of renewable energy. Wind turbines (WTs) usually operate in harsh environments, and hence, are more susceptible to failures compared to other type of energy generators [1]. WT fault diagnosis and prediction are highly important for the wind power industry. WT faults could be caused by exhaustion or manufacturing defect, and they will cause unexpected outage and result in economic loss. One way to minimize this loss is to predict and detect a malfunction before critical damage is done. With this taken into consideration, online methods that can detect fault in real-time have a clear advantage over offline methods. In addition, online diagnosis methods allow automatic remote monitoring of the WT operation conditions, and this is crucial for WT located in remote areas.

WTs are complex systems with multiple subsystems. Each subsystem may be subject to several different types of faults. A comprehensive survey of WT subsystem faults and methods of WT fault diagnosis can be found in [1, 2]. According to [3–6], bearing faults constitute a significant part of failures in WT. The study of this type of failure is not new [3, 7–11]. The most commonly used type of bearings is the ball bearing [1]. Bearing faults can be classified into four groups: inner race, outer race, cage and rollers faults. Many defects start with the fault of bearings.

Different methods are used to identify WT compartments failures, including analyzing mechanical vibrations, acoustic emission, temperature, oil parameters, and electrical signals. Electrical signals, such as current signal, are available remotely, do not require additional sensors, and can be analyzed in real-time [12]. Since faults of a bearing are associated with mechanical defects, they introduce excitations at particular frequencies [7, 13]. For some

bearings, these excitations can be captured by electrical signals [14].

Many bearing fault diagnosis methods have been developed in the literature, such as threshold-based method [13, 15], wavelet-based method [16, 17], and maximum likelihood fault detection [18]. Some existing works model the WT operation states as a Markov process, where the states of the Markov process are used to represent various stages of faults or degradations. The observed data can be used to identify the bearing state [19, 20]. Another category of fault detection methods resort to machine learning or AI-based algorithms such as artificial neural networks (ANN), support vector machine (SVM), decision trees, long short-term memory (LSTM) networks, etc. [21–26]. Both Markov-based and AI-based methods require extensive training before actual detection, and in practice, there might not be sufficient data for training purpose. Fault detection can also be performed with particle filters [27–29]. Particle filters are usually computationally expensive and also require extensive trainings beforehand.

Almost all previous works focus on the accuracy of fault detection and they seldom consider the detection delay. Detection delay is defined as the time difference between the moment that the fault occurs and the time instant that the fault is detected. A small detection delay can result in timely remedial actions, thus prevent catastrophic results due to extended damages and reduce economic loss. Therefore, detection delay is an essential design parameter for WT fault detection or diagnosis.

Quickest or low latency change detection is an online detection method aiming at minimizing the detection delay of an abrupt change in probability distributions of a random process. We propose to develop low latency fault detection method for WTs by analyzing the statistical properties of the stator currents generated by WTs. The stator current can

be modeled as a random process, and the presence of bearing fault will affect the probability distributions of the stator current. The time instant that the fault occurs is a change point, and it can be modeled as a random variable. Generally, change detection methods can be classified into two categories, Bayesian and non-Bayesian (minimax) methods. If the prior probability of the change point is known, then the methods are Bayesian procedures, such as [30, 31]. On the other hand, when the prior probability of the change point is unknown, the low latency change detection methods are developed under the minimax or non-Bayesian criterion. Minimax methods are employed in many practical applications, since it is usually difficult to obtain prior information about change point distribution.

Cumulative sum (CUSUM) procedure [32] and Shiryaev-Roberts (SR) procedure [33] are two most commonly used minimax change detection methods, which aim at minimizing the delay with the worst case change point distribution, under the constraint of a lower bound of average run length (ARL) to false alarm. However, the above-mentioned minimax procedures are developed for systems with single pre-change and single post-change model, and they require the exact information regarding the distribution models before and after the change. Therefore, these procedures cannot be readily applied to WT fault detection since the fault signature might appear at one of several possible frequencies corresponding to several candidates for post-change models. There are limited works with unknown or uncertain post-change models under mimimax criterion. In [34], a quickest change detection algorithm is proposed in order to detect false data injection attacks (FDIA) in smart grids with time-varying dynamic models. An orthogonal matching pursuit CUSUM (OMP-CUSUM) algorithm is proposed to identify the buses under FDIA while minimizing the detection delay

in [35]. However, these two works study the cases where post-change parameter has continuous support, whereas the problem of WT fault detection deals with multiple candidates for post-change models, i.e. the post-change parameter has finite and discrete support.

The objective of this chapter is to develop low latency fault detection methods based on the unique properties of WT bearing faults. Since faults of a bearing introduce excitations (harmonics) into the spectrum of stator current, the analysis is performed in the frequency domain. Statistical analysis show that the amplitude of stator current at a given frequency can be modeled by using Gamma distribution, and the presence of fault will affect the parameters of the Gamma distribution. The theoretical frequency of the excitation caused by a certain fault is determined by the mechanical structure of the bearing. The actual frequency of the fault excitation might deviate from its theoretical value due to the uncertainty of some mechanical parameters. Therefore, in order to detect the fault, the detection algorithm needs to monitor signals on several candidate frequencies. To solve this problem, we propose a multi-candidate low latency fault detection procedure, which is an enhanced version of the CUSUM algorithm. The conventional CUSUM algorithm can only be used to detect change point over a single data stream, yet the proposed multi-candidate detection procedure can be used to detect change point that happens on one of many potential data streams, with the index of the data stream with the change point unknown to the detector. In addition, the theoretical performance bounds on the PFA and ADD of the multi-candidate procedure have been analytically derived. The performance bounds reveal the fundamental tradeoff between detection delay and false alarm. The proposed procedure allows flexible parametric adjustment of the tradeoff between ADD and PFA. Moreover, this detection method does not require a training stage.

The remainder of this chapter is organized as follows. Section 5.3 describes the experimental setup and the process of data collection. The methods for data pre-processing and feature extraction in the frequency domain are presented in 5.4. Section 5.5 studies the statistical behavior of the feature behavior and establishes probability models. The multi-candidate low latency fault detection procedure is proposed and analyzed in Section 5.6. Experiment and simulation results are presented in Section 5.7, and Section 5.8 concludes the chapter.

## 5.3 Experiment Setup and Data Collection

### 5.3.1 Data Collection

Stator current signals obtained from a 160-W Southwest Wind Power Air Breeze direct-drive permanent magnet synchronous generator (PMSG) WT are used in this chapter to test the fault detection algorithm. The procedure for obtaining the stator current signal was described in [13]. The WT used to record the data has six pole pairs ($p = 6$). The data was obtained while the turbine operated in variable wind speed (from 0 to 10 m/s) condition. The shaft of the stator was rotating within the speed range of $6-13$ Hz. The stator current signal was recorded using National Instruments acquisition hardware with a sampling frequency $f_{\text{so}} = 10$ kHz. The signal was recorded within 100-second-long periods every 20 minutes. The total operation time of the WT was about 25 hours. The test bearing (7C55MP4017) was pretreated by removing the lubricant, in order to accelerate degradation. A bearing outer-race fault was generated artificially for a test bearing, as illustrated in [13, Figure 12]. One of the two bearings supporting the shaft was removed to simulate eccentricity. The test wind turbine stopped rotating at the end of the experiment, due to the broken bearing cage.

Although the data were prerecorded and stored offline, they are processed sequentially in an online manner by the faulty detection algorithm to emulate practical application scenarios. That is, during the testing of the fault detection algorithm, the data are fed to the algorithm sequentially in real-time. The algorithm makes decisions based on all current and past data, without any knowledge of the future data that have not been fed to the algorithm. Even though the data were collected over a time period of about 25 hours, experiment results demonstrate that the algorithm can detect wind turbine faults just a few seconds after the occurrence of the fault. Details will be given in Section 5.7.

### 5.3.2 Fault Signatures

Bearing faults introduce excitations on particular frequencies of the stator current. Those excitations are referred to as fault signals. There are four fault modes for bearings: inner race fault, outer race fault, cage fault, and rollers defect. Each one is characterized by a frequency of a fault signal. Denote the frequencies corresponding to inner race fault, outer race fault, cage fault, and rollers defect as $f_{\mathrm{i}}$, $f_{\mathrm{o}}$, $f_{\mathrm{c}}$, and $f_{\mathrm{b}}$, respectively, and they can be calculated as [6]:

$$f_{\mathrm{i}} = 0.5 \cdot N_b \cdot f_r \left(1 + \frac{D_b \cdot \cos\phi}{D_p}\right) \tag{5.1}$$

$$f_{\mathrm{o}} = 0.5 \cdot N_b \cdot f_r \left(1 - \frac{D_b \cdot \cos\phi}{D_p}\right) \tag{5.2}$$

$$f_{\mathrm{c}} = 0.5 \cdot f_r \left(1 - \frac{D_b \cdot \cos\phi}{D_p}\right) \tag{5.3}$$

$$f_{\mathrm{b}} = f_r \cdot \left(\frac{D_b}{D_p}\right) \cdot \left[1 - \left(\frac{D_b \cdot \cos\phi}{D_p}\right)^2\right] \tag{5.4}$$

where $f_r$ is the bearing rotation frequency, $D_b$ is the diameter of rollers, $D_p$ is the rollers pitch diameter, $N_b$ is the number of rollers, and $\phi$ is the rollers' contact angle with races.

Table 5.1: Fault Frequencies

| Faulty compartment | $f_{\text{bf}}$ (Hz) | $f_1 \pm f_{\text{bf}}$ (Hz) |
|---|---|---|
| Inner race | 49.69 | 10.31, 109.69 |
| Outer race | 30.3 | 29.7, 90.3 |
| Cage | 3.78 | 56.22, 63.78 |
| Rollers | 2.28 | 57.72, 62.28 |

The bearing under investigation supports the main shaft, which connects the generator with wind blades. For that reason, the vibrations excited in the bearing affect the stator current by modulating the amplitude of the fundamental frequency $f_1$. The fundamental frequency is defined as the frequency of the main harmonic generated by the wind turbine. Frequencies of fault signals in stator current can be calculated as

$$f_{\text{fault}} = f_1 \pm k \cdot f_{\text{bf}} \tag{5.5}$$

where $f_{\text{bf}} \in \{f_{\text{i}},\ f_{\text{o}},\ f_{\text{b}},\ f_{\text{c}}\}$, and $k = 1, 2, \ldots$ . Throughout this chapter the value $k = 1$ is used.

For the WT that was used to obtain stator current, $f_r = \frac{f_1}{p}$, where $f_1 = 60$ Hz is the fundamental frequency, and $p = 6$ is the number of pole pairs. The parameters of the bearing supporting the shaft are, $D_b = 8$ mm, $D_p = 33$ mm, $N_b = 8$. The contact angle of the rollers is unknown and it is assumed that $\phi = 0$. The frequencies corresponding to the four different types of faults are listed in Table 5.1.

## 5.4 Data Preprocessing and Feature Extraction

The procedures for data preprocessing and feature extraction are described in this section. All procedures described in this section are sequential online algorithms that process new data as they sequentially come in.

### 5.4.1 Synchronous Resampling

The key for fault detection is to extract fault signatures, which are located on particular frequencies relative to fundamental frequency $f_1$. The extraction of fault signatures (5.5) of a signal with non-stationary frequency $f_1$ is rather difficult. The signal-to-noise ratio (SNR) of bearing fault signals is usually small. High frequency resolution is required to reduce the impact of the noise, and more samples for FFT are required. However, the fundamental frequency $f_1$ may drift over time, which leads to time-varying frequencies corresponding to different faults. This makes it difficult to extract the information corresponding to a specific bearing fault.

In order to resolve the problem of fundamental frequency variation, the method of synchronous resampling can be applied [13]. The basic idea is to perform non-uniform sampling such that the phase difference between any two adjacent samples is constant. Details of synchronous resampling can be found in [13]. The resampling frequency used in this chapter is $f_s = 1920$ Hz.

### 5.4.2 Frequency-domain Feature Extraction

After synchronous resampling, the time domain data are converted to the frequency domain via FFT. The resampled time-domain data are divided into length-$N_f$ frames, and FFT is performed on each frame. The time duration for each frame is $T_f = \frac{N_f}{f_s}$ second, and the frequency resolution after FFT is $f_0 = \frac{1}{T_f} = \frac{f_s}{N_f}$. In this chapter, we use $f_s = 1920$ Hz and $N_f = 19200$, and the frequency resolution after FFT is $f_0 = 0.1$ Hz.

Figure 5.1 shows the spectrogram of the data after synchronous resampling. In the spectrogram, we can clearly see the main harmonic of the stator current at 60 Hz. In

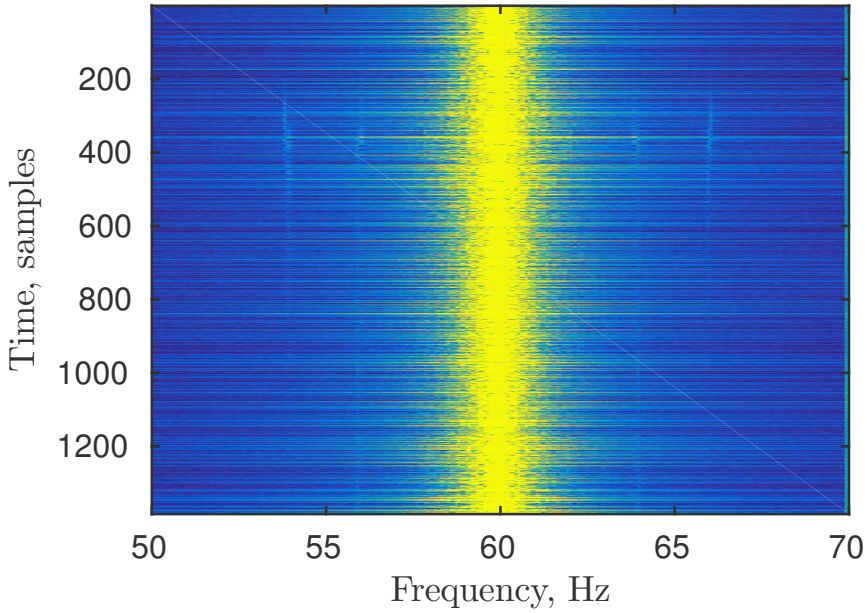


Figure 5.1: FFT of the resampled stator current. Frequency resolution is 0.1 Hz. Excitations due to faults of the bearing are visible on frequencies corresponding to bearing cage and rollers defects.

addition, excitations due to different types of faults are clearly visible around frequencies 54, 56, 58, 62, 64, and 66 Hz.

The fault frequencies shown in the spectrogram do not match with the theoretical values given in Table 1. The mismatch is mainly due to the assumption $\phi = 0$, which is usually non-zero. Nevertheless, the theoretical value gives a rough estimate of the exact location of the fault frequencies. Therefore, during the analysis, we can analyze a range of frequency components centering at the theoretical fault frequencies as $[f_{\text{fault}} - f_w, f_{\text{fault}} + f_w]$, where $f_{\text{fault}}$ is one of the possible fault frequencies, and the window size $2f_w$ is chosen to be an integer multiple of the frequency resolution $f_0$.

For a given fault frequency $f_{\text{fault}}$, we can extract the amplitude of the signals over the block of frequencies in $[f_{\text{fault}} - f_w, f_{\text{fault}} + f_w]$ and form them as a feature vector. For the $n$-th

FFT frame, denote the signal amplitude vector over the frequency range $[f_{\text{fault}}-f_w, f_{\text{fault}}+f_w]$ as $\mathbf{s}_n = [s_{n1}, s_{n2}, \cdots, s_{nw}]$, where $s_{nj}$ is the signal amplitude at frequency $f_{\text{fault}}-f_w+(j-1)f_0$ in the $n$-th frame, and $w = 2\frac{f_w}{f_0} + 1$ is the number of elements in the feature vector. Since the fault feature could appear at any one of the frequencies in $\{f_{\text{fault}} - f_w + (j-1)f_0\}_{j=1}^{w}$, we denote the set of frequencies as candidate frequencies. In this chapter, $f_w = 0.5$ Hz is used, which results in a feature vector size $w = 11$.

### 5.4.3 Energy Normalization

Due to the fluctuation of the operation environment, such as wind speed, the power of the feature vector changes with respect to time. Energy normalization is performed to compensate for power fluctuations.

The energy normalization is performed by using a sliding window approach. The energy of each sample is normalized by the average energy of samples in the current and the previous $N_w$ frames. The normalization of the $j$-th element in the $n$-th frame can be expressed as

$$x_{nj} = \frac{s_{nj}}{\sqrt{e_{nj}}} \tag{5.6}$$

where

$$e_{nj} = \frac{1}{N_w + 1} \sum_{k=n-N_w}^{n} |s_{kj}|^2 \tag{5.7}$$

is the average energy of the current and past $N_w$ samples on the $j$-th candidate element.

With the normalized samples, we can form the normalized feature vector as $\mathbf{x}_n = [x_{n1}, x_{n2}, \cdots, x_{nw}]^T$.

The normalized feature vector $\mathbf{x}_n$ is a random vector. We will detect the presence of fault by analyzing the statistical properties of the components in $\mathbf{x}_n$.

## 5.5 Statistical Analysis of Feature Vector

Statistical analysis is performed over the frequency domain feature vector to identify their statistical properties with and without different types of faults.

### 5.5.1 Statistical Distribution

We obtain histogram by using the experiment data corresponding to the samples on the $j$-th candidate frequency $\{x_{nj}\}_n$. It is found that the distribution of $x_{nj}$ can be accurately modeled by using Gamma distribution with probability density function (PDF) given as

$$p(x;\alpha,\beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-x\beta}}{\Gamma(\alpha)}, \tag{5.8}$$

where $\Gamma(\alpha)$ is the Gamma function, and $\alpha$ and $\beta$ are the parameters of the Gamma distribution. The parameters $\alpha$ and $\beta$ depend on the element index $j$ as well as the presence of fault.

The mean, $m_x$, and variance, $\sigma_x^2$, of a Gamma distributed random variable $x$ are

$$m_x = \frac{\alpha}{\beta}, \quad \sigma_x^2 = \frac{\alpha}{\beta^2} \tag{5.9}$$

### 5.5.2 Online Parameter Estimation

We propose to estimate the distribution parameters, $\alpha$ and $\beta$, online by using a sliding window approach. At each frame, the parameters for signals on one candidate frequency can be estimated by using the corresponding data from the current and the past $N_w$ frames.

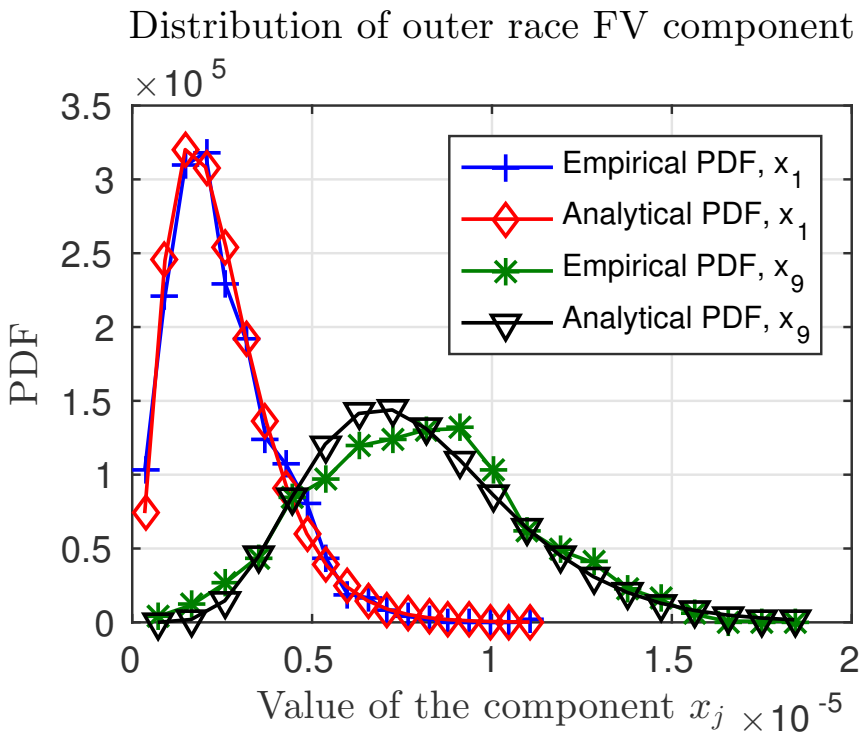


Figure 5.2: Empirical and analytical Gamma distribution of outer race feature vector components.

The mean and variance of the $j$-th element at the $n$-th frame can be estimated as

$$\hat{m}_{nj} = \frac{1}{N_w + 1} \sum_{k=n-N_w}^{n} x_{kj} \tag{5.10}$$

$$\hat{\sigma}^2_{nj} = \frac{1}{N_w} \sum_{k=n-N_w}^{n} (x_{kj} - \hat{m}_{nj})^2 \tag{5.11}$$

It should be noted that the factor $\frac{1}{N_w}$ is used in (5.11) to ensure the estimation is unbiased.

From (5.9), at the $n$-th frame, we can obtain an estimate of the parameters $\alpha$ and $\beta$ for the $j$-th element as

$$\hat{\alpha}_{nj} = \frac{\hat{m}^2_{nj}}{\hat{\sigma}^2_{nj}}, \quad \hat{\beta}_{nj} = \frac{\hat{m}_{nj}}{\hat{\sigma}^2_{nj}} \tag{5.12}$$

Figure 5.2 shows the empirical and analytical pdfs of $x_{n1}$ and $x_{n9}$ from the feature vector related to the outer race fault, by using parameters estimated with (5.12). In the experiment

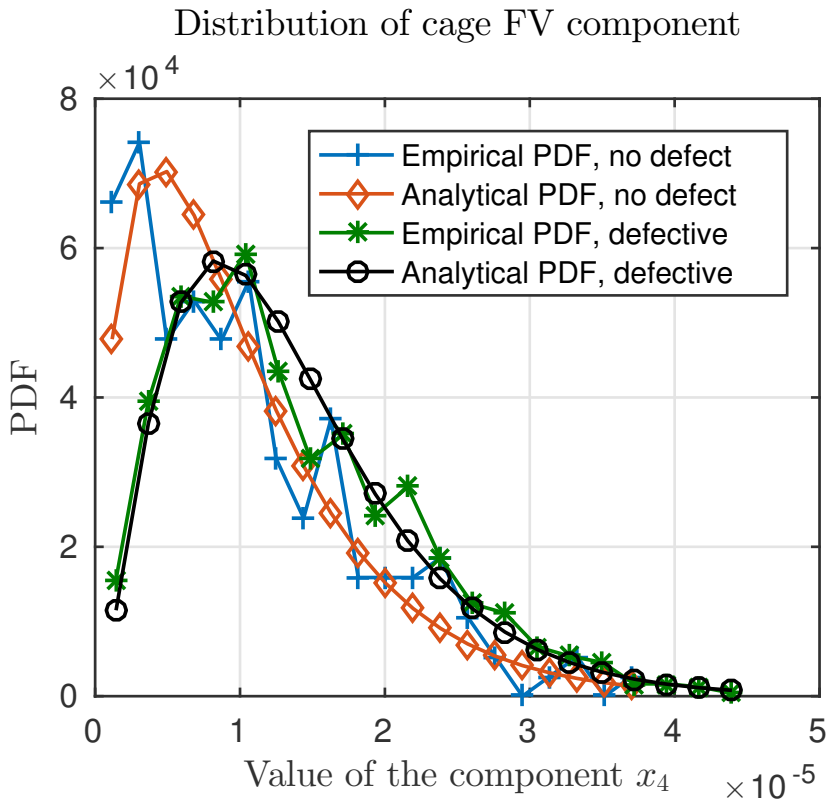


Figure 5.3: Empirical and analytical Gamma distribution of cage feature vector components in normal and defective conditions.

setup, the outer race fault is artificially introduced at the beginning of the experiment and its signature is shown in $x_{n9}$, which is at the frequency of 30.0 Hz. There is no fault signature at $x_{n1}$, which is at the frequency of 29.2 Hz. The analytical pdf is obtained by evaluating (5.8) by using parameters estimated from (5.12). The empirical pdf is obtained by normalizing the histogram of the experiment data. There is a good match between the analytical and empirical pdf. Also, there is an obvious difference in distribution between $x_{n1}$ and $x_{n9}$ due to the presence of fault signature on $x_{n9}$.

Figure 5.3 shows the distribution of $x_{n4}$ from the feature vector related to the cage fault, before and after the fault occurs. In the experiment, the cage fault starts to appear at $n = 130$ frame, and the fault signature is at $56.2 - f_w + (j-1)f_0 = 56$ Hz, which corresponds

to an index of $j = 4$ in the feature vector. The pdf corresponding to the non-defective case is calculated by using data with $n < 130$, and the pdf corresponding to the defective case is calculated by using data with $n \geq 130$. There is a clear difference between the distributions before and after the fault occurs.

Moreover, when the fault is not present, it is discovered that the elements in a fault feature vector are identically distributed. That is, when the fault is not present, $\hat{\alpha}_{nj} \approx \hat{\alpha}_{mj}$ and $\hat{\beta}_{nj} \approx \hat{\beta}_{mj}$ for $m \neq n$. Such a property can be used to further improve the estimation accuracy of the parameters.

It should be noted the proposed frequency domain modeling is robust against noise. Noise in real-world systems can be modeled as additive white Gaussian noise (AWGN), which has uniform power spectrum density in the frequency domain. The flat noise spectrum means the noise power is uniformly distributed over the entire frequency bands, which results in a very low noise power over a small frequency range. On the other hand, the fault signature is embedded over a very small frequency range. Thus we can significantly increase the signal-to-noise ratio (SNR) by focusing on the analysis of the signal over a very small range of frequency bands. As a result, the presence of noise will have negligible impacts regarding the performance of proposed algorithm.

### 5.6 Low Latency Fault Detection

The sequential low latency fault detection algorithm is presented in this section. The detection algorithm is formulated by exploring the statistical distributions of the feature vector before and after the fault occurs.

### 5.6.1 Formulation of Hypothesis

For the fault detection, define the two hypotheses for the variable $x_{nj}$

$$H_0: \quad \text{No fault} \tag{5.13}$$

$$H_1: \quad \text{Fault is present} \tag{5.14}$$

As shown in Figs. 5.2 and 5.3, the presence of fault will affect the values of $\alpha$ and $\beta$ for the corresponding element in the feature vector. Denote the parameters before and after the change point as $\left(\alpha^{(0)}, \beta^{(0)}\right)$ and $\left(\alpha^{(1)}, \beta^{(1)}\right)$, respectively.

For a Gamma distributed random variable $x$ with energy normalized to unit, from (5.9), we have

$$m_x^2 + \sigma_x^2 = \frac{\alpha^2}{\beta^2} + \frac{\alpha}{\beta^2} = 1 \tag{5.15}$$

which yields

$$\beta = \sqrt{\alpha(\alpha+1)} \tag{5.16}$$

Therefore we only need to characterize the impacts of fault on $\alpha$, and the corresponding $\beta$ can be inferred from (5.16).

To quantify the impacts of fault, define a new parameter

$$\alpha_\Delta = \frac{\alpha^{(1)}}{\alpha^{(0)}} \tag{5.17}$$

For the outer race fault shown in Figure 5.2, the average value of $\alpha_\Delta$ is 1.4. For the cage fault shown in Figure 5.3, the average value of $\alpha_\Delta$ is 2.9.

In practical system design, $\alpha_\Delta$ is unknown beforehand. We will treat $\alpha_\Delta$ as an adjustable design parameter that can be set by using prior knowledge. We will show the impact of the choice of $\alpha_\Delta$ on the detection performance in the section of experiment results.

With the definition of $\alpha_\Delta$ and the estimation of $\hat{\alpha}_j^{(0)}$ using (5.12), the hypotheses can be formulated as

$$H_0 : x_{nj} \sim \text{Gamma}\left(\hat{\alpha}_j^{(0)}, \hat{\beta}_j^{(0)}\right) \tag{5.18}$$

$$H_1 : x_{nj} \sim \text{Gamma}\left(\hat{\alpha}_j^{(1)}, \hat{\beta}_j^{(1)}\right) \tag{5.19}$$

where

$$\hat{\alpha}_j^{(1)} = \hat{\alpha}_j^{(0)} \alpha_\Delta \tag{5.20}$$

$$\hat{\beta}_j^{(k)} = \sqrt{\hat{\alpha}_j^{(k)}(\hat{\alpha}_j^{(k)} + 1)}, \text{ for } k = 0, 1 \tag{5.21}$$

With the hypothesis formulated in (5.18), define the likelihood ratio (LR) as

$$\lambda_{nj} = \frac{p(x_{nj}; \hat{\alpha}_j^{(1)}, \hat{\beta}_j^{(1)})}{p(x_{nj}; \hat{\alpha}_j^{(0)}, \hat{\beta}_j^{(0)})} \tag{5.22}$$

Combining (5.8) and (5.22) yields

$$\begin{aligned} \log \lambda_{nj} = \hat{\alpha}_j^{(1)} \log \hat{\beta}_j^{(1)} - \hat{\alpha}_j^{(0)} \log \hat{\beta}_j^{(0)} + \left(\hat{\alpha}_j^{(1)} - \hat{\alpha}_j^{(0)}\right) \log x_{nj} - x_{nj} \left(\hat{\beta}_j^{(1)} - \hat{\beta}_j^{(0)}\right) \\ + \log \Gamma(\hat{\alpha}_j^{(0)}) - \log \Gamma(\hat{\alpha}_j^{(1)}) \end{aligned} \tag{5.23}$$

### 5.6.2 Low Latency Fault Detection Algorithm

Based on the hypothesis and LR formulated in the previous subsection, we propose a multi-candidate low latency fault detection algorithm.

Assume a certain fault occurs at frame $\theta$ and the fault signature is on the $\mu$-th candidate frequency in the feature vector. Both $\theta$ and $\mu$ are random variables. That is, when $n < \theta$, $x_{nj}$ follows the distribution specified by $H_0$; when $n \geq \theta$, $x_{n\mu}$ follows the distribution specified by $H_1$, and $x_{nk}$ follows the distribution of $H_0$ for all $n$ and $k \neq \mu$.

Denote the detected change point as $\hat{\theta}$. Then the PFA can be defined as

$$\text{PFA} = \mathbf{P}(\hat{\theta} < \theta) = \sum_{k=1}^{\infty} \pi_k \mathbf{P}(\hat{\theta} < k) \tag{5.24}$$

where $\pi_k = \mathbf{P}(\theta = k)$ is the prior probability of the change point.

The well known CUSUM and SR procedures are designed for the change point detection in a single time sequence. They cannot be readily applied to the problem encountered in this chapter, where there are multiple parallel data streams (multiple candidate frequencies), and the change point occurs in at most one of the data stream (candidate frequency).

To address this problem, we propose a multi-candidate low latency detection procedure as follows.

*Definition 5.1*: (Multi-candidate detection procedure) With the frequency domain feature vector $\mathbf{x}_n$, the detected change point is

$$\hat{\theta} = \inf\{n : C_n \geq A\} \tag{5.25}$$

where we set $\inf\{\varnothing\} = \infty$, $A$ is a pre-defined threshold determined by the PFA, the test statistic $C_n$ is defined as $C_n = \sum_{j=1}^{w} C_{n,j}$, $C_{n,j}$ is the test statistic for the $j$-th element, and it can be recursively calculated as

$$C_{n+1,j} = \max(1, C_{n,j})\lambda_{n+1,j} \tag{5.26}$$

and $C_{1,j} = 1$.

This detection procedure is an enhanced version of the CUSUM procedure. It combines the test statistic $\{C_{n,j}\}_{j=1}^{w}$ to determine the change point. The CUSUM procedure can be considered as a special case of the multi-candidate procedure when $w = 1$. It should be noted that the multi-candidate test procedure proposed in Definition 5.1 is different from the procedure proposed in [36]. In [36], the summation is performed over the logarithm

of the CUSUM statistic, i.e., $\sum_{j=1}^{w} \log C_{n,j}$, which corresponds to the multiplication of the linear CUSUM statistic $C_{n,j}$.

### 5.6.3 Performance Analysis

Next, we study the performance of the multi-candidate test procedure. Let $\mathbf{P}_k$ and $\mathbb{E}_k$ denote the probability measure and the corresponding expectation when the change occurs at $\theta = k$. With such a notation, $\mathbf{P}_\infty$ and $\mathbb{E}_\infty$ can be used to represent the probability measure and expectation before the change point, that is, $\theta = \infty$.

*Lemma 5.1*: The test statistic $C_n$ is a sub-martingale under the probability measure $\mathbf{P}_\infty$, and $\mathbb{E}_\infty[C_n] \leq nw$.

*Proof:* To show that $C_n$ is a sub-martingale under $\mathbf{P}_\infty$, it is sufficient to prove that $\mathbb{E}_\infty[C_{n+1}|C_n] \geq C_n$. For the LR $\lambda_{n,j}$, we have

$$\mathbb{E}_\infty(\lambda_{n,j}) = \int \frac{p(x_{nj}; \hat{\alpha}_j^{(1)}, \hat{\beta}_j^{(1)})}{p(x_{nj}; \hat{\alpha}_j^{(0)}, \hat{\beta}_j^{(0)})} p(x_{nj}; \hat{\alpha}_j^{(0)}, \hat{\beta}_j^{(0)}) dx_{nj} = 1$$

Based on the definition of $C_n$, we have

$$\begin{aligned}
\mathbb{E}_\infty[C_{n+1}|C_n] &= \sum_{j=1}^{w} \mathbb{E}_\infty[C_{n+1,j}|C_n] \\
&= \sum_{j=1}^{w} \max(1, C_{n,j}) \mathbb{E}_\infty(\lambda_{n+1,j}) \\
&\geq \sum_{j=1}^{w} C_{n,j} = C_n
\end{aligned}$$

Thus $C_n$ is a sub-martingale.

In addition,

$$\mathbb{E}_\infty(C_n) = \mathbb{E}_\infty \left( \sum_{j=1}^{w} \max_{1 \leq k \leq n} \Lambda_{k:n,j} \right) \leq \mathbb{E}_\infty \left( \sum_{j=1}^{w} \sum_{k=1}^{n} (\Lambda_{k:n,j}) \right) = nw.$$

■

*Theorem 5.1*: Assume the elements in the feature vector $\mathbf{x}_n$ are independent. With the multi-candidate procedure defined in 5.1, the probability of false alarm is upper bounded by

$$\text{PFA} \leq \frac{\bar{\theta} w}{A} \tag{5.27}$$

where $\bar{\theta} = \sum_{k=0}^{\infty} \pi_k k$ is the prior mean of the change point $\theta$, and $w$ is the number of candidates.

*Proof:*

$C_n$ is a sub-martingale under $\mathbf{P}_\infty$. Based on Doob's inequality, we have

$$\mathbf{P}_\infty(\hat{\theta} < n) = \mathbf{P}_\infty \left( \max_{1 \leq k < n} C_n \geq A \right) \leq \frac{\mathbb{E}_\infty(C_n)}{A} \leq \frac{nw}{A} \tag{5.28}$$

Therefore, the PFA can be calculated as

$$\text{PFA} = \sum_{k=1}^{\infty} \pi_k \mathbf{P}_k(\hat{\theta} < k) = \sum_{k=1}^{\infty} \pi_\infty \mathbf{P}_k(\hat{\theta} < k) \leq \sum_{k=1}^{\infty} \pi_k \frac{kw}{A} = \frac{\bar{\theta} w}{A}.$$

■

Next, we study the detection delay of the multi-candidate test procedure. It is difficult to obtain the exact ADD of the detection procedure. Instead, we will obtain an asymptotic upper bound. To facilitate analysis, define

$$f(x) \underset{x \to x_0}{\preceq} g(x) \iff \lim_{x \to x_0} \frac{f(x)}{g(x)} \leq 1 \tag{5.29}$$

*Theorem 5.2*: Assume the elements in the feature vector $\mathbf{x}_n$ are independent, and the frames are independent in time. Let PFA $< \alpha$. As $\alpha \to 0$, we have

$$\mathbb{E}[\hat{\theta} - \theta | \hat{\theta} > \theta] \underset{\alpha \to 0}{\preceq} \frac{1}{1-\alpha} \frac{|\log \alpha| + \log \bar{\theta} + \log w}{D_{10}} \tag{5.30}$$

where $\bar{\theta}$ is the priori mean of the change point, $w$ is the length of the feature vector, and

$$D_{10} = \int p(x; \alpha^{(1)}, \beta^{(1)}) \log \frac{p(x; \alpha^{(1)}, \beta^{(1)})}{p(x; \alpha^{(0)}, \beta^{(0)})} dx \tag{5.31}$$

is the Kullback-Leibler (KL) divergence between the post-change distribution $p(x; \alpha^{(1)}, \beta^{(1)})$ and the pre-change distribution $p(x; \alpha^{(0)}, \beta^{(0)})$.

*Proof:* Assume the fault happens at the $k$-th frame on the $j$-th element. Define a new stopping time

$$\tau(A) = \inf\{n \geq k : Z_j^{k:n} \geq \log A\} \tag{5.32}$$

where

$$Z_j^{k:n} = \sum_{l=k}^{n} \log \lambda_{l,j} \tag{5.33}$$

From (5.26), we have

$$C_{n,j} = \max_{1 \leq k \leq n} \prod_{l=k}^{n} \lambda_{l,j} \tag{5.34}$$

Thus $C_n \geq C_{n,j} \geq \prod_{l=k}^{n} \lambda_{l,j}$ when $n \geq k$, or equivalently, $\log C_n \geq Z_j^{k:n}$. As a result, we have $\hat{\theta} \leq \tau(A)$.

Based on the strong law of large numbers, under the probability measure $\mathbf{P}_k$, $\frac{1}{n} Z_j^{k:k+n-1}$ almost surely converges in probability $\mathbf{P}_k$ to the KL divergence $D_{10} = \mathbb{E}_k[\log \lambda_{l,j}]$,

$$\frac{1}{n} Z_j^{k:k+n-1} \xrightarrow[n \to \infty]{\mathbf{P}_k - a.s.} D_{10}. \tag{5.35}$$

Define

$$T_k = \sup \left\{ n \geq 1 : \left| \frac{1}{n} Z_j^{k:k+n-1} - D_{10} \right| > \epsilon \right\}. \tag{5.36}$$

If $\tau(A) - k > T_k$, then from (5.36) we have

$$\left| \frac{1}{\tau(A) - k} Z_j^{k:\tau(A)-1} - D_{10} \right| \leq \epsilon, \quad \text{if } \tau(A) - k > T_k \tag{5.37}$$

which implies

$$\tau(A) - k \leq \frac{Z_j^{k:\tau(A)-1}}{D_{10} - \epsilon}, \quad \text{if } \tau(A) - k > T_k \tag{5.38}$$

Based on the definition of $\tau(A)$ in (5.32), we have

$$Z_j^{k:\tau(A)-1} < \log A \tag{5.39}$$

Combining (5.38) and (5.39) results in

$$\tau(A) - k \leq \frac{\log A}{D_{10} - \epsilon}, \quad \text{if } \tau(A) - k > T_k \tag{5.40}$$

When $\log(A) > 0$ and $\epsilon < D_{10}$, the following inequality is true for both $\tau(A) - k > T_k$ and $\tau(A) - k \leq T_k$

$$\tau(A) - k \leq \frac{\log A}{D_{10} - \epsilon} + T_k, \quad \text{if } \log A > 0 \text{ and } \epsilon < D_{10} \tag{5.41}$$

Given the convergence condition in (5.35), we have $\mathbb{E}(T_k) < \infty$. Since $\epsilon$ can be arbitrarily small, we can let $\epsilon \to 0$. From Theorem 5.1, we can set $A = \frac{\bar{\theta} w}{\alpha}$ to guarantee PFA $< \alpha$. Thus when $\alpha \to 0$ and $\epsilon \to 0$,

$$\mathbb{E}_k[\tau(A) - k] \underset{\alpha \to 0}{\preceq} \frac{\log|\alpha| + \log\bar{\theta} + \log w}{D_{10}}. \tag{5.42}$$

Since $\hat{\theta}$ is upper bounded by $\tau(A)$, we have

$$\mathbb{E}_k[\hat{\theta} - k] \underset{\alpha \to 0}{\preceq} \frac{\log|\alpha| + \log\bar{\theta} + \log w}{D_{10}}. \tag{5.43}$$

When $\alpha \to 0$, the right hand side of (5.43) is positive. Thus the inequality in (5.43) still holds if we replace $\hat{\theta}_1 - k$ by $(\hat{\theta}_1 - k)^+$.

The ADD can be calculated as

$$\mathbb{E}[\hat{\theta} - \theta | \hat{\theta} > \theta] = \frac{1}{\mathbf{P}_\infty(\hat{\theta} \geq \theta)} \sum_{k=1}^{\infty} \pi_k \mathbb{E}_k(\hat{\theta} - k)^+ \tag{5.44}$$

With the constraint PFA $< \alpha$, we have $\mathbf{P}_\infty(\hat{\theta} \geq \theta) = 1 - \text{PFA} \geq 1 - \alpha$. Combining (5.32), (5.44), and the above results yields (5.30).

■

In Theorem 5.2, the asymptotic upper bound of the ADD is inversely proportional to the KL divergence $D_{10}$. The KL divergence measures the difference between two distributions. When the difference between the two distributions before and after the change point is large, the detection delay is in general smaller. The asymptotic ADD upper bound is a decreasing function of the PFA $\alpha$, which demonstrates the tradeoff relationship between the PFA and ADD. In addition, the asymptotic ADD upper bound is proportional to $\log w$, which means we can reduce the detection delay by reducing the number of candidate frequencies. On the other hand, the value of $w$ must be large enough to ensure that the frequency corresponding to the faulty feature is included in the feature vector.

### 5.6.4 Complexity Analysis

In order to apply the proposed multi-candidate detection procedure to on-line data in each time-frame, the procedure needs to go through 3 stages: 1) data preprocessing, 2) feature extraction, and 3) test statistic computation. In stage 1, the computation is dominated by the fast Fourier transform (FFT) operation that has a complexity of $\mathcal{O}(N_f \log N_f)$, where $N_f$ is the frame size. The second stage involves operations such as energy normalization, mean and variance estimation. Each of these operations has a complexity of $\mathcal{O}(N_w)$, where $N_w$ is the number of past time-frames (samples) used for parameter estimation and energy normalization. Finally, the third stage is dominated by the operation described in equation (5.26), which has a complexity of $\mathcal{O}(w)$, where $w$ is the feature vector size, i.e., the number of candidate frequencies.

In this chapter, we set $N_f = 19200, N_w = 100$, and $w = 11$. As a result, the overall complexity of the proposed algorithm is dominated by stage 1, which is $\mathcal{O}(N_f \log N_f)$. That

means the frame size $N_f$ is the most dominating factor in the computational complexity of the algorithm. A frame of $N_f = 19200$ samples with a sampling rate of $f_s = 1920$ Hz corresponds to a frame duration of $T_f = 10$ seconds. Thus as long as the algorithm can finish the computation within 10 seconds, the algorithm can be implemented in an online fashion. Our simulation and experiment results show that the calculation time of the algorithm is much less than 10 seconds, thus the algorithm can be directly applied to online data.

## 5.7 Numerical Analysis

In this section, we demonstrate the performance of the proposed bearing fault detection algorithm with both experiment and synthesized simulation data. Since the experiment data is obtained from only one trial, it is important to obtain the results with simulation data to analyze the bounds on performance metrics.

### 5.7.1 Experiment Results

We test the performance of the proposed bearing fault detection algorithm with the experiment data. In the data processing, we set the FFT frame size as $N_f = 19,200$, a resampling frequency of $f_s = 1,920$ Hz, which result in a frequency resolution of $f_0 = 0.1$ Hz and a frame duration of $T_f = \frac{N_f}{f_s} = 10$ sec. The size of the feature vector is $w = 11$, which corresponds to a frequency span of $wf_0 = 1.1$ Hz. The number of past time-frames (samples) used for parameter estimation and energy normalization is set as $N_w = 100$.

In the experiment setup, the inner race fault and outer race fault are artificially introduced at the beginning of the experiment. The cage fault and roller fault starts to appear at the $n = 130$th frame.

The performance of the proposed low latency fault detection algorithm can be tuned by setting the values of two parameters, the degree of change $\alpha_\Delta$ and the detection threshold $A$. The value of $\alpha_\Delta$ is determined by the distributions before and after the change point. If the location of the change point is known, we can estimate the value of $\alpha_\Delta$ by using data before and after the change point. The value $\alpha_\Delta$ may be different for different types of faults. For example, for the outer race fault shown in Figure 5.2, the average value of $\alpha_\Delta$ is 1.4, while for the cage fault shown in Figure 5.3, the average value of $\alpha_\Delta$ is 2.9. We obtained these empirical values using offline analysis based on the experimental data, where we knew the exact location of the change point, i.e., the time instant when a particular fault occurs.

However, in case of online change point (fault) detection, we will not have the exact knowledge of $\alpha_\Delta$ beforehand, as it requires the exact location of change point. Therefore, in online change detection, $\alpha_\Delta$ is treated as an adjustable design parameter that can be set by using prior knowledge. The performance of the proposed algorithm under different values of $\alpha_\Delta$ is shown next.

Since the experiment data are obtained under one trial, the PFA cannot be directly calculated from the experiment data. To address this issue, we propose to study the behaviors of false alarm by using the frequency-domain data streams without fault. Specifically, for each of the 4 faults, we extract features from several groups of frequency-domain data centered around frequencies which are outside the range of $[f_{\text{fault}} - f_w, f_{\text{fault}} + f_w]$, where $f_{\text{fault}}$ denotes the theoretical fault frequency (c.f. Table 1). The detection procedure is then applied to these frequency-domain data streams and the total number of false alarms is recorded. The frequency-domain false alarm ratio is defined as the ratio between the total number of false alarms and the total number of time-frames multiplied by the number of

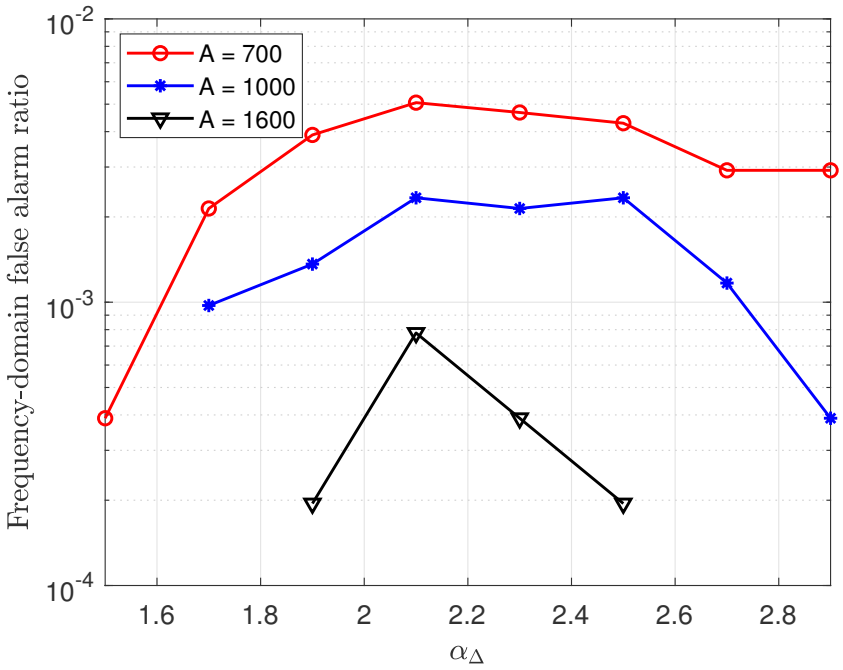


Figure 5.4: Frequency-domain false alarm ratio v.s. $\alpha_\Delta$.

groups. The frequency-domain false alarm ratio can serve as an indicator for PFA. Figure 5.4 shows the frequency-domain false alarm ratio as a function of $\alpha_\Delta$, under various values of the threshold $A$. The relationship between frequency-domain false alarm ratio and $\alpha_\Delta$ is not monotonic. Under a fixed threshold $A$, the frequency-domain false alarm ratio is concave or quasi-concave in $\alpha_\Delta$. The choice of $\alpha_\Delta$ needs to consider the tradeoff between the detection delay and frequency-domain false alarm ratio.

The detection delay of the proposed multi-candidate detection algorithm is shown in Figure 5.5. This metric is obtained by running the proposed detection procedure 4 times with four different sets of candidate frequencies corresponding to the 4 types of faults and then computing the average delay in detecting these 4 faults. The detection delays are shown in the units of frames, with the duration of each frame being 10 seconds. Under different configurations, the detection delay ranges between 27 to 48 frames, which correspond to an

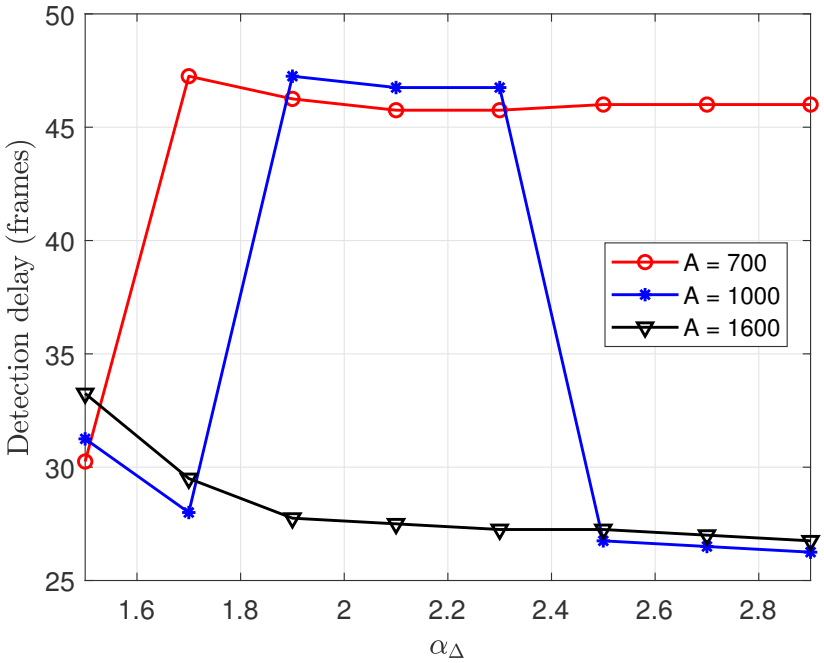


Figure 5.5: Detection delay v.s. $\alpha_\Delta$.

absolute delay of 270 to 480 seconds. Based on our statistical analysis of the experiment data, the average true value of $\alpha_\Delta$ is around 2.

Figure 5.6 shows the detection delay as a function of the frequency domain false alarm ratio. This figure demonstrates the tradeoff relationship between delay and false alarm. For comparison, results obtained from three other detection algorithms, namely the conventional impulse detection algorithm [13], SUM-Shiryaev-Roberts (SUM-SR) algorithm [37], and CUSUM-GLRT [38] algorithm, are also shown in the figure. SUM-SR procedure is an enhanced version of the SR procedure, where all the local SR statistics corresponding to individual post-change models are combined together. In all algorithms, we set $\alpha_\Delta = 2.0$, and different tradeoff points are achieved by adjusting the threshold value $A$. It is observed that the proposed multi-candidate algorithm slightly outperforms the existing algorithms.

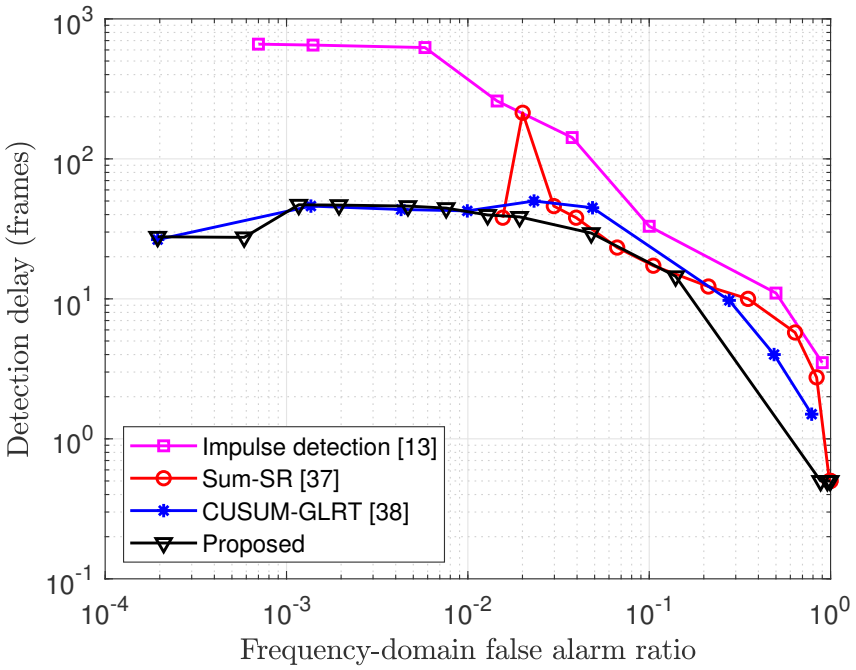


Figure 5.6: Detection delay v.s. frequency-domain false alarm ratio.

### 5.7.2 Simulation Results

Synthesized simulation data are used to verify the analytical bounds derived in this chapter. In the simulation setup, we model the occurrence of the fault as a geometric distribution, that is, $\pi_k = (1-\rho)^{k-1}\rho$, where $\rho = 0.1$. The value of $\alpha_\Delta$ is set to 2. The pre-change distribution parameters for each candidate frequency $\alpha_j^{(0)}$ follows a uniform distribution between $[1 \times 10^{-5}, 5 \times 10^{-5}]$. The simulation results are obtained by averaging over 10,000 Monte-Carlo trials, where only a single type of fault is simulated in each trial.

Figure 5.7 shows the PFA as a function of the detection threshold $A$. Under the same threshold $A$, the CUSUM-GLRT algorithm has the lowest PFA, followed by the proposed algorithm, and the PFA of the SUM-SR algorithm is significantly higher than the other two algorithms. The PFA of all three algorithms are lower than the analytical upper bounds developed in Theorem 1. It should be noted that a lower PFA does not necessarily mean a

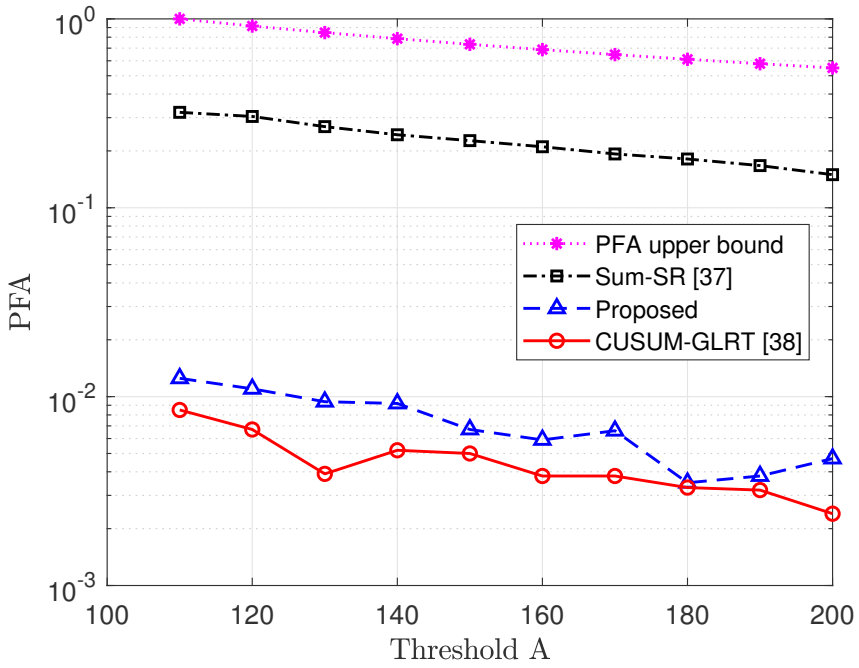


Figure 5.7: Probability of false alarm v.s. threshold.

better performance due to the effects of the ADD.

The ADDs of the three different algorithms are shown in Figure 5.8 as a function of the detection threshold. Under the same threshold $A$, the SUM-SR algorithm has the lowest ADD at the cost of a large PFA, followed by the proposed multi-candidate detection procedure and CUSUM-GLRT, respectively. The ADD of the proposed low latency algorithm is lower than the asymptotic upper bound presented in Theorem 5.2.

The tradeoff relationship between ADD and PFA for the simulated data is illustrated in Figure 5.9, where the ADD is shown as a function of PFA. Different points on the tradeoff curve are obtained by adjusting the detection threshold $A$. Among the four algorithms, the proposed algorithm achieves the best ADD-PFA tradeoff, followed by the SUM-SR algorithm, the CUSUM-GLRT algorithm, and the impulse detection algorithm, respectively. Since the impulse detection algorithm is not a quickest change detection algorithm, the ADD of the

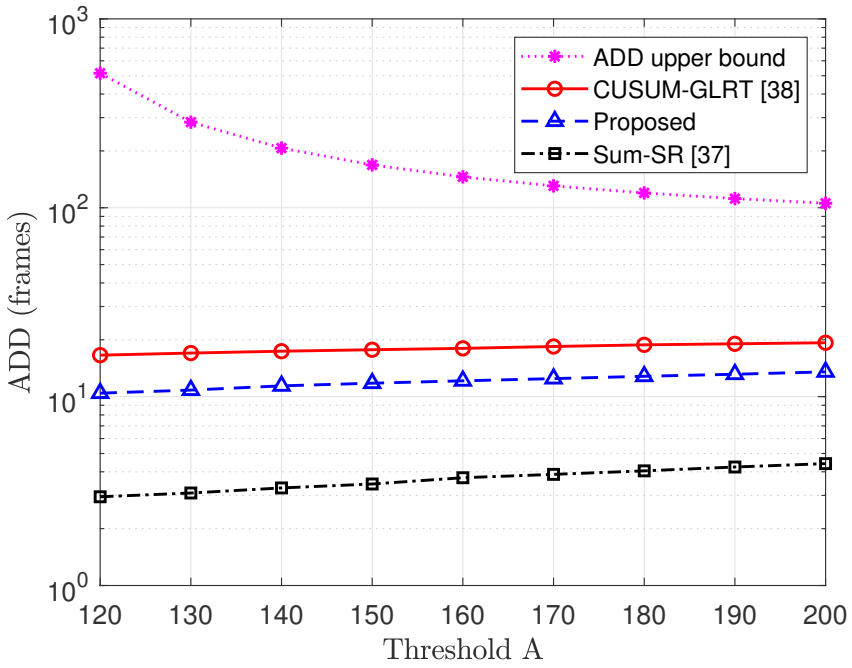


Figure 5.8: Average detection delay v.s. threshold.

impulse detection algorithm is significantly larger than the other three algorithms, which are developed by using sequential statistics. At a PFA of 0.04, the ADDs of the proposed algorithm, the SUM-SR algorithm, the CUSUM-GLRT algorithm, and the impulse detection algorithm are 7, 7.5, 10.7, and 382 frames, respectively.

## 5.8 Conclusion

The detection of bearing faults of direct-drive wind turbines have been studied in this chapter under the framework of low latency change point detection. The amplitude of stator current at a given frequency was modeled by using the Gamma distribution, and the presence of fault will affect the parameters of the Gamma distribution. We defined a new parameter, $\alpha_\Delta$, to measure the impact of the fault on the Gamma distribution. A multi-candidate low latency change point detection algorithm, which includes the conventional CUSUM algorithm as a

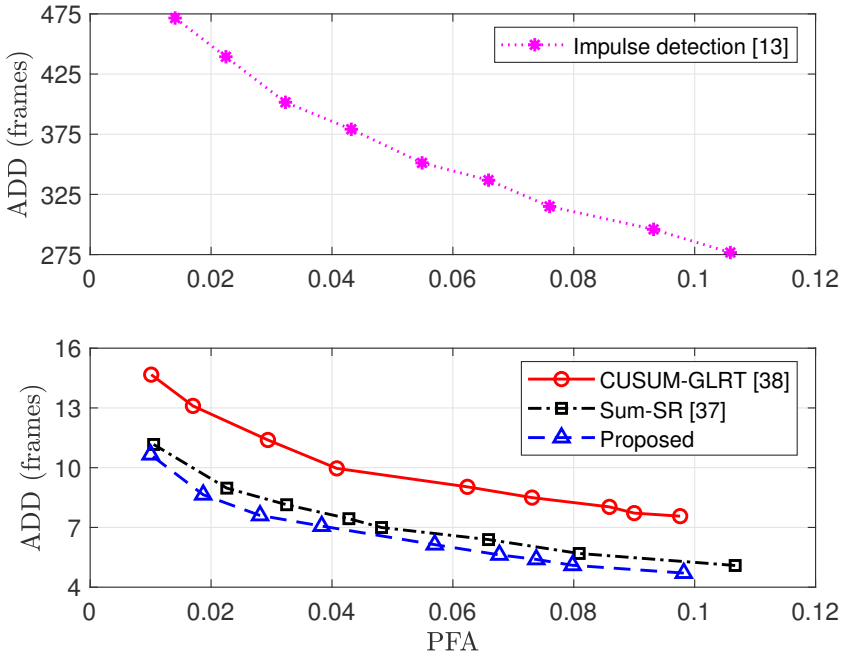


Figure 5.9: Average detection delay vs probability of false alarm.

special case, has been proposed to detect the various faults. The theoretical performance of the proposed algorithm has been analytically identified in terms of upper bounds of the probability of false alarm and average detection delay. This algorithm can flexibly achieve different tradeoff between the PFA and ADD. This proposed low latency fault detection algorithm does not require a training phase and can be readily applied to the collected data in real-time. Experiment and simulation results demonstrated that the proposed algorithm can detect faults in real-time with minimum delays, and it outperforms existing algorithms. The proposed multi-candidate change detection can also be applied to other change detection scenarios with the change occurring in one of many possible data streams.

## References


[1] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis - part i: Components and subsystems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6536–6545, 2015.

[2] ——, "A survey on wind turbine condition monitoring and fault diagnosis - part ii: Signals and signal processing methods," 2015.

[3] M. Blödt, P. Granjon, B. Raison, and G. Rostaing, "Models for bearing damage detection in induction motors using stator current monitoring," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 4, pp. 1813–1822, 2008.

[4] B. Lu, Y. Li, X. Wu, and Z. Yang, "A review of recent advances in wind turbine condition monitoring and fault diagnosis," in *Power Electronics and Machines in Wind Applications (PEMWA)*. IEEE, 2009, pp. 1–7.

[5] Z. Daneshi-Far, G. Capolino, and H. Henao, "Review of failures and condition monitoring in wind turbine generators," in *XIX International Conference on Electrical Machines (ICEM)*. IEEE, 2010, pp. 1–6.

[6] X. Gong and W. Qiao, "Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3419–3428, 2013.

[7] R. R. Schoen, T. G. Habetler, F. Kamran, and R. Bartfield, "Motor bearing damage detection using stator current monitoring," *IEEE Transactions on Industry Applications*, vol. 31, no. 6, pp. 1274–1279, 1995.

[8] F. Immovilli, M. Cocconcelli, A. Bellini, and R. Rubini, "Detection of generalized-roughness bearing fault by spectral-kurtosis energy of vibration or current signals," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4710–4717, 2009.

[9] B. M. Ebrahimi, J. Faiz, and M. J. Roshtkhari, "Static-, dynamic-, and mixed-eccentricity fault diagnoses in permanent-magnet synchronous motors," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4727–4739, 2009.

[10] W. Yang, P. J. Tavner, and M. Wilkinson, "Condition monitoring and fault diagnosis of a wind turbine synchronous generator drive train," *IET Renewable Power Generation*, vol. 3, no. 1, pp. 1–11, 2009.

[11] F. Immovilli, A. Bellini, R. Rubini, and C. Tassoni, "Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison," *IEEE Transactions on Industry Applications*, vol. 46, no. 4, pp. 1350–1359, 2010.

[12] L. He, L. Hao, D. Pan, and W. Qiao, "Detection of single-axis pitch bearing defect in a wind turbine using electrical signature analysis," in *IEEE International Electric Machines Drives Conference (IEMDC)*, San Diego, CA, USA, May 2019, pp. 31–36.

[13] X. Gong and W. Qiao, "Current-based mechanical fault detection for direct-drive wind turbines via synchronous sampling and impulse detection," *IEEE Transactions on Industrial Electronicss*, vol. 62, no. 3, pp. 1693–1702, 2015.

[14] ——, "Current-based eccentricity detection for direct-drive wind turbines via synchronous sampling," in *Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2013, pp. 2972–2976.

[15] C. Peeters, P. Guillaume, and J. Helsen, "Vibration-based bearing fault detection for operations and maintenance cost reduction in wind energy," *Renewable Energy*, vol. 116, pp. 74–87, 2018.

[16] X. Gong, W. Qiao, and W. Zhou, "Incipient bearing fault detection via wind generator stator current and wavelet filter," in *36th Annual Conference on IEEE Industrial Electronics Society (IECON)*. IEEE, 2010, pp. 2615–2620.

[17] W. Teng, W. Wang, H. Ma, Y. Liu, Z. Ma, and H. Mu, "Adaptive fault detection of the bearing in wind turbine generators using parameterless empirical wavelet transform and margin factor," *Journal of Vibration and Control*, vol. 25, no. 6, pp. 1263–1278, 2019.

[18] I. S. Bozchalooi and M. Liang, "Parameter-free bearing fault detection based on maximum likelihood estimation and differentiation," *Measurement Science and Technology*, vol. 20, no. 6, p. 065102, 2009.

[19] H. Ocak and K. A. Loparo, "A new bearing fault detection and diagnosis scheme based on hidden markov modeling of vibration signals," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5. IEEE, 2001, pp. 3141–3144.

[20] X. Zhang, R. Xu, C. Kwan, S. Y. Liang, Q. Xie, and L. Haynes, "An integrated approach to bearing fault diagnostics and prognostics," in *Proceedings of the American Control Conference*. IEEE, 2005, pp. 2750–2755.

[21] X. An, D. Jiang, and S. Li, "Application of back propagation neural network to fault diagnosis of direct-drive wind turbine," in *World Non-Grid-Connected Wind Power and Energy Conference (WNWEC), 2010*. IEEE, 2010, pp. 1–5.

[22] Y. Wang and D. Infield, "Neural network modelling with autoregressive inputs for wind turbine condition monitoring," in *International Conference on Sustainable Power Generation and Supply (SUPERGEN)*. IET, 2012, pp. 1–6.

[23] M.-S. An, S.-J. Park, J.-S. Shin, H.-Y. Lim, and D.-S. Kang, "Implementation of automatic failure diagnosis for wind turbine monitoring system based on neural network," in *Multimedia and Ubiquitous Engineering*. Springer, 2013, pp. 1181–1188.

[24] I. Abdallah, V. Dertimanis, H. Mylonas, K. Tatsis, E. Chatzi, N. Dervilis, K. Worden, and E. Maguire, "Fault diagnosis of wind turbine structures using decision tree learning algorithms with big data," *Safety and Reliability–Safe Societies in a Changing World*, pp. 3053–3061, 2018.

[25] W. Teng, H. Cheng, X. Ding, Y. Liu, Z. Ma, and H. Mu, "Dnn-based approach for fault detection in a direct drive wind turbine," *IET Renewable Power Generation*, vol. 12, no. 10, pp. 1164–1171, 2018.

[26] J. Lei, C. Liu, and D. Jiang, "Fault diagnosis of wind turbine based on long short-term memory networks," *Renewable energy*, vol. 133, pp. 422–432, 2019.

[27] B. Zhang, C. Sconyers, C. Byington, R. Patrick, M. Orchard, and G. Vachtsevanos, "Anomaly detection: A robust approach to detection of unanticipated faults," in *International Conference on Prognostics and Health Management (PHM)*. IEEE, 2008, pp. 1–8.

[28] M. E. Orchard and G. J. Vachtsevanos, "A particle-filtering approach for on-line fault diagnosis and failure prognosis," *Transactions of the Institute of Measurement and Control*, vol. 31, no. 3-4, pp. 221–246, 2009.

[29] B. Zhang, C. Sconyers, C. Byington, R. Patrick, M. Orchard, and G. Vachtsevanos, "A probabilistic fault detection approach: application to bearing fault detection," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, 2011.

[30] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 22–46, 1963.

[31] S. Nath and J. Wu, "Bayesian quickest change point detection with multiple candidates of post-change models," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Anaheim, CA, USA, Nov. 2018.

[32] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.

[33] S. W. Roberts, "A comparison of some control chart procedures," *Technometrics*, vol. 8, pp. 411–430, 1966.

[34] S. Nath, I. Akingeneye, J. Wu, and Z. Han, "Quickest detection of false data injection attacks in smart grid with dynamic models," *IEEE Journal of Emerging and Selected Topics in Power Electronics (Early Access)*, Aug. 2019. [Online]. Available: https://doi.org/10.1109/JESTPE.2019.2936587

[35] I. Akingeneye and J. Wu, "Low latency detection of sparse false data injections in smart grids," *IEEE Access*, vol. 6, pp. 58 564–58 573, Oct. 2018.

[36] Y. Mei, "Efficient scalable schemes for monitoring a large number of data streams," *Biometrika*, vol. 97, no. 2, pp. 419–433, 2010.

[37] A. G. Tartakovsky and V. V. Veeravalli, "Change-point detection in multichannel and distributed systems," *Applied Sequential Methodologies: Real-World Examples with Data Analysis*, vol. 173, pp. 339–370, 2004.

[38] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, 1998.

# Chapter 6

# Conclusions

This chapter summarizes the main contributions of this dissertation and lists some possible directions for future research.

## 6.1 Contributions

The content of this dissertation focuses on the design of algorithms that enhance the operation stability and security in various systems with imperfect post-change models by detecting any type of abnormality in the system as quickly as possible. The algorithms are built under the quickest change detection (QCD) framework which is appropriate for low latency anomaly detection. The main findings of the dissertation are summarized as follows.

First, QCD algorithms for detecting changes in systems with multiple post-change models are designed under both Bayesian and non-Bayesian settings. The theoretical ADDs of the proposed algorithms have been obtained through asymptotic analysis and it is proven that the proposed algorithms are asymptotically optimal in terms of detection delay. Simulation results have shown that the proposed algorithms outperform the existing algorithms in terms of detection delay under similar PFA constraints.

Second, a QCD algorithm is studied in order to detect false data injection attacks (FDIA) in smart grids with time-varying dynamic models. To distinguish between FDIA and sudden system change, a time-varying dynamic model is used to represent the dynamic state transitions. A new normalized Rao-CUSUM detector is developed by analyzing the statistical

properties of dynamic state estimations, such that the detector minimizes the worst-case detection delay while accurately distinguishing FDIA from sudden system changes. Unlike existing algorithms that rely on measurement correlation to discriminate false data from sudden system changes, the proposed algorithm can detect any false data including those injected into correlated measurements. Simulation results demonstrate that the proposed algorithm can accurately detect and remove false data injections or system faults with minimum delays, thus improving the security and resilience of smart grids.

Third, a QCD algorithm was designed for the detection of bearing faults of direct-drive wind turbines. After analyzing the statistical properties of the stator currents in the frequency domain, it has been discovered that the current amplitude at a given frequency can be modeled by using the Gamma distribution, and the presence of fault will affect the parameters of the Gamma distribution. A multi-candidate quickest change point detection algorithm, which is an enhanced version of CUSUM algorithm, has been proposed to detect the various faults. The theoretical performance of the proposed multi-candidate detection procedure has been analytically identified in terms of upper bounds of PFA and ADD. Through analysis with both the experiment data and simulation data, it is observed that proposed algorithm outperforms other existing algorithms.

In summary, the research in this dissertation covered developing theoretical algorithms for low latency anomaly detection in systems where the information regarding the post-change model is unknown or imperfect. Analytical results and simulation results are obtained to validate the performance of the proposed algorithms. In addition, two practical applications are studied which can utilize the low latency anomaly detection algorithms in real-time. One of the applications is the detection of false data injection attacks in smart grids with dynamic

models and the other one is the bearing fault detection of direct-drive wind turbines using stator current.

## 6.2 Future Work

We now discuss possible future directions in which the topics discussed in this dissertation can be extended.

• *Non-i.i.d. settings:* A major assumption that has been used throughout this dissertation is the i.i.d. assumption, i.e., the observations are identically and independently distributed. The QCD problem is studied under non-i.i.d. settings in [1, 2], where the post-change models are assumed to be known. It would be interesting to study the QCD problem with both the constraints of imperfect models and non-i.i.d. settings.

• *Change isolation:* The quickest change detection problem can also be studied in combination with the change isolation problem [3, 4]. In this problem, the objective is not only to detect the change, but also to isolate the change, i.e., to identify the possible post-change distribution at the time of detection. It would be interesting to investigate the joint problem of change detection and isolation in systems with imperfect post-change models. For example, in the case of our study in Chapter 5, it will be more beneficial if we are able to not only detect the bearing faults quickly, but also to identify which particular fault (inner race, outer race, cage, or roller) has occurred.

• *Machine learning approach:* In this dissertation, all the QCD algorithms that have been presented are based on the theory of sequential hypothesis testing. However, in recent years, there has been a rise of machine learning-based approaches in the field of change detection. Change detection methods using kernel-based non-parametric statistics is one of the ways to

tackle the issue of imperfect post-change model, since these methods make fewer assumptions on the distributions than traditional parametric approaches [5,6]. In [6], it is discovered that one-class support vector machine (SVM) can be trained online with unlabeled data sets because of its unsupervised nature, which breaks the strong assumption that most of the existing methods are based on. Additionally, by utilizing the predictive power of Long Short Term Memory (LSTM) network, methods are developed in [7,8] to detect anomalies in time series data. It is demonstrated that by modeling the normal behavior of a time series through LSTM networks, deviations from normal behavior can be detected without any prespecified context window or preprocessing.

To summarize, all the above-mentioned approaches combine the problem of change detection with various machine learning techniques. This provides an impetus to study further on this topic with a specific focus on low latency detection without requiring complete model information.

## References


[1] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, Nov. 1998.

[2] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic bayesian theory of quickest change detection," *Theory of Probability & Its Applications*, vol. 49, no. 3, pp. 458–497, 2005.

[3] A. G. Tartakovsky, "Multidecision quickest change-point detection: Previous achievements and open problems," *Sequential Analysis*, vol. 27, no. 2, pp. 201–231, 2008.

[4] T. Banerjee, Y. C. Chen, A. D. Dominguez-Garcia, and V. V. Veeravalli, "Power system line outage detection and identification-a quickest change detection approach," in *IEEE*

*International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3450–3454.

[5] F. Desobry, M. Davy, and C. Doncarli, "An online kernel change detection algorithm," *IEEE Trans. Signal Processing*, vol. 53, no. 8-2, pp. 2961–2974, 2005.

[6] Z. Zhang and H. Shen, "Online training of svms for real-time intrusion detection," in *18th International Conference on Advanced Information Networking and Applications, 2004. AINA 2004.*, vol. 1. IEEE, 2004, pp. 568–573.

[7] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "Long short term memory networks for anomaly detection in time series," in *Proceedings*. Presses universitaires de Louvain, 2015, p. 89.

[8] L. Bontemps, J. McDermott, N.-A. Le-Khac *et al.*, "Collective anomaly detection based on long short-term memory recurrent neural networks," in *International Conference on Future Data and Security Engineering*. Springer, 2016, pp. 141–152.